Award Number: W81XWH-13-1-0080

TITLE: "Psycho-Motor and Error Enabled Simulations: Modeling Vulnerable Skills in the Pre-Mastery Phase - Medical Practice Initiative Procedural Skill Decay and Maintenance (MPI-PSD)"

PRINCIPAL INVESTIGATOR: Carla Pugh, MD, PhD

CONTRACTING ORGANIZATION: University of Wisconsin System
Madison, WI 53715-1218

REPORT DATE: December 2016

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE December 2016	**2. REPORT TYPE** Final	**3. DATES COVERED** 13Mar2013-12Sep2016

4. TITLE AND SUBTITLE
"Psycho-Motor and Error Enabled Simulations: Modeling Vulnerable Skills in the Pre-Mastery Phase - Medical Practice Initiative Procedural Skill Decay and Maintenance (MPI-PSD)"

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-13-1-0080

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S) Carla Pugh, MD, PhD; Doug Wiegmann, PhD; Caprice Greenberg, MD, PhD; Ferdinando Mussa-Ivaldi, PhD; Eugene “Chip” Foley, MD; Felix Huang, PhD; Andrea Mason, PhD; Rob Radwin, PhD; Thomas Yen, PhD; Shlomi Laufer, PhD; Rebecca Ray, PhD, Anne-Lise D’Angelo, MD; Eran Gwillim, MD; Calvin Kwan, BS; Drew Rutherford, MS; Elaine Cohen, Med; Katherine Law, BS; Shannon DiMarco, BA; Donna Hankins, MS; Caitlin Jenewein, BS, Jay Nathwani, MD, Hossein Mohamadipanah, PhD

email: pugh@surgery.wisc.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Wisconsin System
21 N. Park Street, Suite 6401,
Madison, WI 53715-1218

8. PERFORMING ORGANIZATION REPORT

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR’S ACRONYM(S)

11. SPONSOR/MONITOR’S NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This term of the grant has been comprised of three main directions. The first primary direction focused on simulation preparation, personnel training, and conducting data collection sessions. Based on recommendations identified from pilot simulations, necessary modifications were made to each of the procedural stations. Next, the entire team traveled to the institutions involved in the study, set up all the simulation stations, interacted with participants, and collected data to bring back to the lab for analysis. Additionally, there was a focus on collecting the same data set from surgeons who had retired from the field at varying levels. The second direction of effort focused on the organization, database and transcription coding, and analysis of this data. Analysis was expedited using participant workbooks, common error checklists, and video recordings made during participant data collections. Additional data such as images of final products were obtained and coded. The third direction focused on the dissemination of findings learned in the initial analyses performed. A set of scores validating our data was developed and will be used for future coding, analysis and reporting methods to participants. Multiple abstracts and posters were created for surgical conferences attended. These works concentrated on data from pre and post participation surveys, perceptions of skill reduction, motor control in robotic tasks, longitudinal comparisons and comparisons between robotic and simulation tasks. We incorporated lessons learned from our years of work into our research extension and have developed a system to analyze the data collected.

15. SUBJECT TERMS Refined development of simulated procedure stations; Training of personnel in human subjects research for data collection; Completion of data collection from participating institutions; Development and execution of Performance Review Tool; Organization, coding, and transcribing of collected data; Analysis of qualitative survey and quantitative procedure simulation data; Dissemination of results at surgical conferences

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	97	USAMRMC
					19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Introduction

The entirety of the grant primarily focused on identifying key differentiating performance factors in the pre-mastery and mastery phases. Identifying these factors was primarily done by evaluating survey results from three years of data collections, and analyzing performance on the six stations including cross analysis between stations. The lab was able to compare performance for study participants at one station to those whose data was analyzed from previously recorded in both 2014 and 2015 along with simulation based LVH repairs. Station design was updated and optimized to allow for the best possible data collection.

Data was also collected from surgeons who varied in retirement. This performance data was captured to best understand how expert surgeons retain their knowledge after deciding to stop a regular surgical rotation.

Summary for Entire Research Effort

For review, the following four objectives guided this work:

Objective One: To evaluate mental rehearsal as an intervention for skill decay in the pre-mastery phase.

Objective Two: To identify key differentiating performance factors for the pre-mastery and mastery phases.

Objective Three: To develop a generalizable, multi-variable, predictive model of skills decay.

Objective Four: To develop an efficient and effective set of assessment tools and individualized training recommendations to counteract skills decay.

We collected data from participants at nine Midwestern academic institutions over the summers of 2014, 2015 and 2016. These individuals were compensated for their participation and were consented that they were allowed to opt of the study at any time. Prior to data collection, we suited all participants to best capture their data. This included pairing them with an audio recorder, individual video capturing systems and individual motion sensors to collect their movements as they proceeded through four simulation stations and two virtual reality, psycho-motor function tests. Participant workbooks were also distributed.

Site	Number of Participants for Data Collection Period			
	Summer 2014	Summer 2015	Summer 2016	Retired Surgeons
UW-Madison	13	11	12	2
Loyola	5	1	0	0
Mayo	12	11	10	0
UIC	2	0	0	0
University of Chicago	9	8	5	1
Rush University	3	0	3	0
Northwestern	3	3	0	1
Washington University	0	7	11	0
MCW	0	5	5	0
Other	0	0	0	10
All Sites	**47**	**46**	**46**	**14**

Table 1: Demographic information for data collections.

Data collection took approximately 2.5 hours and varied depending on participant skill level and speed. All participants were give the same amount of time to complete a simulated station (typically 15 minutes) and several simulated stations also contained cognitive scenarios which were asked after performing a station.

Figure 1: Spring 2014 Participant being suited up with motion tracking for data collection. The original set up of the participants was quickly redesigned to include better security for the tracking sensors.

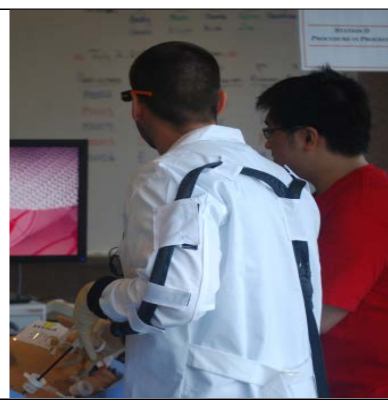
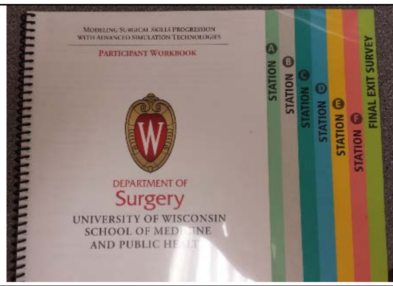


Figure 2: Summer 2014 Participant completing the simulated laparoscopic ventral hernia repair. The updated protocol for participants was to wear video recording glasses and a coat which held the sensor wires securely.

Figure 3: Summer 2015 Participant completing the simulated injured bowel repair. The final protocol for participants was to wear video recording headset to minimize slippage that occurred with the glasses.

Participation Workbook



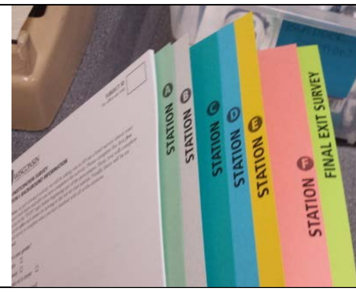
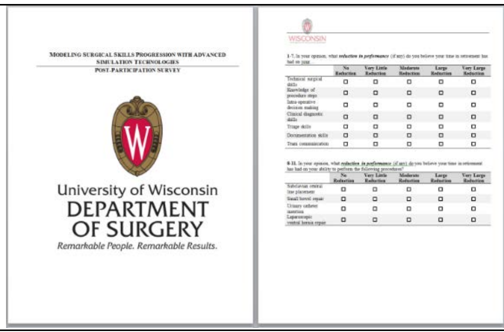


Figure 4 and 5: The workbook for residents that participated in the study. While minor updates were made year to year, the style and format of these remained consistent throughout the study.



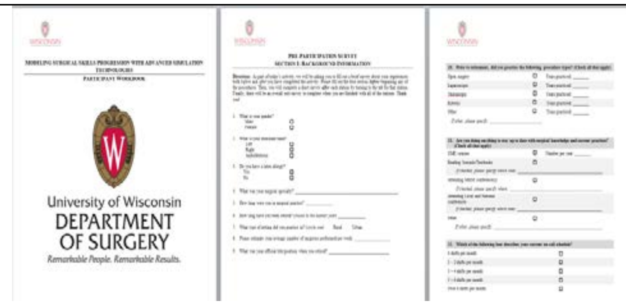


Figure 6 and 7: The workbook for retired surgeons that participated in the study.

All participants were required to complete a pre-participation survey prior to beginning any simulated procedures. This survey began with a brief demographic section followed by current post-graduate year (PGY). PGY was intended to be indicative of clinical experience, however an addition was made in 2015 to this section to collect procedure-specific information on clinical experience. Information was also recorded on how often the participant is on-call or moonlights, as well as how many and what types of procedures are performed during these shifts. Next, data was collected about the participant's perceived skills decay, confidence, and difficulty for both general and procedure-specific skills. Finally, participant's specified the frequency at which they experience a number of conditions, such as headache and nausea, that can be experienced upon exposure to virtual reality environments. After participation in each simulated and virtual reality scenario, residents were asked to complete the same procedure-specific confidence and difficulty questions and record any relevant comments.

Retired surgeons were asked similar questions; only modified to understand their previous and current clinical obligations (if applicable).

The following are publications containing analyses from the participation workbook followed by a brief description of the discovery made in the publication:

- D'Angelo, A. L., Ray, R. D., Jenewein, C. G., Jones, G. F., & Pugh, C. M. (2015). Residents' perception of skill decay during dedicated research time. *Journal of Surgical Research, 199*(1), 23-31.

  This study looked at resident perception of skill reduction during dedicated research time. It was hypothesized that residents would perceive greater potential skill reduction during research time for clinical abilities and procedures that are more complex and require increased decision making. Residents perceived a greater reduction in bowel anastomosis ($M=3.0$, $SD=0.96$) and LVH repair ($M=2.84$, $SD=1.03$) skills compared to subclavian central line insertion ($M=2.13$, $SD=0.81$) and urinary catheterization ($M=1.50$, $SD=0.51$) ($F(1,37)=36.59$, $p<.001$). The more months residents had already spent in the lab correlated with greater perceived reduction in technical surgical skills ($r(37)=.40$, $p=.007$) and knowledge of procedure steps ($r(37)=.36$, $p=.015$). Overall, the residents in this study expected moderate skills decay in open-loop tasks (bowel anastomosis and LVH repair), in technical surgical skills, and knowledge of procedure steps. Simulation-based assessments and training may provide a mechanism for maintaining skills and keeping confidence grounded in experience during dedicated research time.

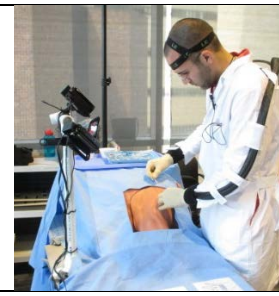
- Jones, G. F., Law, K. E., Jenewein, C. G., Ray, R. D., DiMarco, S. M., & Pugh, C. M. (2016). Assessment during clinical years changes resident perception of skills decay. *The American Journal of Surgery.* (Accepted)

  In this study, residents were hypothesized to have a different perception of their skill decay depending on whether or not they had previously been assessed using simulation. Their pre- and post- surveys for four simulated clinical procedures were evaluated for perceived skill decay, confidence, and difficulty of the procedure. Overall, residents perceived skill reduction in both procedure specific and general surgical skills. Returning residents reported larger skill reductions in urinary catheterization ($t(22.8)=2.96$, $p=.007$) and in technical surgical skills ($t(35)=2.37$, $p=.023$). New residents reported more confidence in performing the entire urinary catheterization procedure ($F(1,42) = 7.55$, $p = .009$) and both groups of residents reported decreases in overall confidence post-simulation ($F(1,42) = 4.53$, $p = .039$). New residents also increased in confidence for the LVH repair after performing the procedure, while returning resident confidence either remained the same or decreased slightly ($F(1, 43)=3.60$, $p=.065$). Finally, both new and returning residents rated the LVH repair as the most difficult procedure, followed by small bowel repair, subclavian central line, and urinary catheterization pre-procedure. The urinary catheterization procedure showed the largest difference between the groups in perceived difficulty ($F(1,43)=4.14$, $p=.019$). In conclusion, differences between new and returning residents in confidence, perceived skill decay and procedural difficulty reveal vulnerable skills that can be targeted by simulated interventions.

Work in Progress:

In the future, resident on-call and moonlighting shift frequency will be analyzed for correlations to clinical performance data (*Objective Two*). In addition, perceived skills-decay, confidence, and difficulty ratings both pre- and post-simulation performance will be longitudinally analyzed through the three years of the study along with the comparison to the retired surgeons (*Objective Three*). This data will reveal any trends in perceived skill reduction throughout years of minimal to no clinical practice (*Objective One*). Longitudinal trends related to skill reduction will be compared to the longitudinal clinical performance data (*Objective One and Objective Three*). Finally, a training and assessment tool will be developed to minimize and test the amount of skills decay in research residents (*Objective Four*).

Subclavian Central Line Placement

		
Figure 7: Summer 2014 setup of central line placement scenario. General feedback we received from participants was that the simulation was not realistic to what a real life situation would present.	*Figure 8: Summer 2015 setup of central line placement scenario. Major changes were made based on feedback including creating a surface to correctly place the central line kit as well as masking the majority of the simulator and only exposing the necessary "skin".*	*Figure 9: Summer 2015 completion of the central line simulation station. This figure also shows the additional cameras added to the protocol to ensure better data gathering of catheter insertion.*

An immersive simulated scenario was created to assess resident decision making and technical ability in the placement of a subclavian central line. In data collections held in 2014, 2015, and 2016 of this study, residents were presented with a clinical scenario of a 67 year old male who presents in septic shock with a blood pressure of 90/40, heart rate of 110 and temperature of 39°C. The simulated patient was already placed in a Trendelenburg position, prepped and draped, and injected with local anesthetic. Participants were provided with a subclavian central line kit containing all necessary instruments including a needle, syringe, guidewire, scalpel, dilator and catheter. During subclavian central line placement, motion metric data was collected from the participant's hands, and video and audio recordings were taken. A researcher was tasked with recording any technical errors made by the participant during the simulation for further review. The participant was expected to confirm correct placement of the catheter by aspirating through the brown catheter port upon completion. Following central line placement in the CentraLine Man trainer, each participant was presented with cognitive scenarios and asked to detail complications and how to address them in one of two difficult clinical scenarios and two of four misplaced central line x-ray scenarios. These responses were audio recorded and later transcribed for analysis of decision making ability. Finally, the participant completed a brief survey where they were asked to indicate their confidence in their performance, as well as their perceived difficulty of the task.

The following are publications containing analyses from this simulation followed by a brief description of the discovery made in the publication:

- Nathwani, J. N., Fiers, R. M., Ray, R. D., Witt, A. K., Law, K. E., DiMarco, S., & Pugh, C. M. (2016). Relationship between technical errors and decision-making skills in the junior resident. *Journal of Surgical Education, 73*(6), e84-e90.

The purpose of this study was to co-evaluate resident technical errors and decision-making capabilities during placement of a subclavian CVC in a simulated procedure. It was hypothesized that there would be significant correlations between scenario based decision making skills and technical proficiency. It was also hypothesized that residents would have some problems anticipating common difficulties and generating solutions associated with line placement in cognitive scenarios. Participants averaged a greater than expected 1.87 errors ($SD=1.5$, $t(44)=3.82$, $p<.001$). The number of total errors performed during the subclavian central line procedure correlated negatively with the number of possible difficulties that residents could list while placing a central line in the given clinical scenario ($r(33)=-.419$, $p=.021$). The number of total errors showed a significant negative correlation with the number of top solutions that residents could generate in response to anticipated difficulties in placing a subclavian central line ($r(33)=-.383$, $p=.044$). With central line insertion listed as a Phase 1 basic/core skill by the resident skills curriculum, our presumption was that the majority of residents would have minimal to no technical errors or decision-making difficulties. The unexpected findings on resident subclavian central line placement indicates that further exposure to subclavian CVC using simulation could translate into improved performance. Once competency is achieved, simulation could still be employed to help maintain skills.

- Mohamadipanah, H., Parthiban, C., Nathwani, J., Rutherford, D., DiMarco, S., & Pugh, C. (2016). Can a virtual reality assessment of fine motor skill predict successful central line insertion? *The American Journal of Surgery, 212*(4), 573-578.

  The aim of this study was to identify a relationship between the ability of surgical residents to discriminate forces in a VR environment and the use of specific motor movements during needle insertion for subclavian CVC. It was hypothesized that ability in fine motor discrimination, as assessed in a VR environment, would correlate significantly with the gross motor movements during subclavian CVC. It was found that the metric "correct discrimination" had a negative correlation ($r=-3802$, $P=.0119$), and the metric "wrong discrimination" had a positive correlation ($r=.3317$, $P=.0298$) to the "insertion time" metric. This means that persons with better fine motor skills had a shorter needle insertion time. In addition, the metric "correct discrimination" had a negative correlation ($r=-4628$, $P=.0018$) to the "idle time" metric, and the metric "wrong discrimination" had a positive correlation ($r=.4108$, $P=.0062$) to the "idle time" metric". In essence, those persons with excellent fine motor force discrimination have less pauses during needle insertions, and those with poor fine motor force discrimination have longer pauses. These findings indicate that haptic sensation is an important motor skill that could affect performance of needle insertion in subclavian CVC. Overall, VR training and assessments could be used to guide fine motor skills training for residents.

- Witt, A. K., Nathwani, J. N., DiMarco, S. M., Pugh, C. M. (2017). Skill loss and maintenance for research residents: A longitudinal study. *Society of Black Academic Surgeons.* Abstract Submitted.

In this study, it was hypothesized that a longitudinal simulated central line assessment would reveal no progression in surgical skills in junior research residents. Of the residents who participated in this assessment twice, 14.3% made the same error twice and 53.4% of participants made new errors in their second year. Performance significantly improved in only two out of twelve (16.7%) categories on the common error checklist in second year performance ($p<0.05$). The large portion of residents making the same or new errors in their second year is suggestive that adequate skill maintenance is not occurring. The lack of improvement compounded by the commitment of new errors shows there is measurable skills decay. These findings support the use of focused assessments and feedback to promote skill maintenance during residents' laboratory years.

Work in progress:

Analysis of longitudinal performance data is underway for two out of the three years of the study. The third year of longitudinal performance data will be added within the next 3 months as data coding and transcription is completed. A comprehensive, longitudinal analysis of decision-making ability will be analyzed using resident responses to the presented cognitive scenarios. These longitudinal decision-making trends will then be compared to the trends in clinical performance of the subclavian central line placement using the coded error data (*Objectives Two and Three*).

Bowel Anastomosis

		
Figure 10: Summer 2014 setup of injured bowel anastomosis scenario. General feedback we received from participants was that the simulation was not realistic to what a real life situation would present.	*Figure 11: Summer 2015 setup of Injured bowel anastomosis scenario. Major changes were made based on feedback including creating a shallower surface to complete the procedure.*	*Figure 12: Summer 2015 completion of the injured bowel anastomosis station. This figure also shows the researcher acting as a surgical assistant which was available to participants all years of the study.*

The bowel anastomosis station was designed with intent of being challenging to participants. Participants were presented with a clinical vignette: a 28-year-old male presents as a trauma after suffering from multiple gunshot wounds. Primary and secondary survey reveal retention of one bullet in the mid-abdomen. The patient is taken for exploratory laparotomy where a portion of small intestine has two enterotomies. To simulate the open environment, a portion of animal cadaver intestine is placed on a tray in front of the participant. Two separate injuries are created for the participant to identify: one injury is approximately 1 cm in diameter whereas the second injury is approximately 0.5 cm in diameter, the injuries are spaced by approximately 1 cm. The participant is then tasked with repair of the bowel with basic open surgery tools, excluding staplers, and a variety of absorbable and non-absorbable sutures to perform the bowel repair. Participants are provided with an operative assistant at the level of medical student. Multiple streams of data were collected including checklist, reflective of cognitive and technical ability, video, motion, discourse, and final product. Post collection protocols included, coding video performance, transcription of discourse, and video-analysis of final product. After completion of the procedure, a post-performance survey was performed exploring confidence in resident performance. Data collection occurred for three consecutive years. Differences in data collection were two fold and occurred as changes to improve protocol from the 2014 to 2015 year. First, an additional camera was placed in order to have a clearer view on participant hand movements allowing greater analyzation of hand performance and suturing technique and patterns. Secondly, in order to understand if participants had strong knowledge of how to use their operative assistant, the bowel clips were taken away originally used to affix the bowel in place.

The following are publications containing analyses from this simulation followed by a brief description of the discovery made in the publication:

- Gannon, S. J., Law, K. E., Ray, R. D., Nathwani, J. N., DiMarco, S. M., D'Angelo, A. L. D., & Pugh, C. M. (2016). Do resident's leadership skills relate to ratings of technical skill? *Journal of Surgical Research*, *206*(2), 466-471.

  This study sought to compare general surgery research residents' survey information regarding self-efficacy ratings to their observed performance during a simulated small bowel repair. Their observed performance ratings were based on their leadership skills in directing their assistant. An analysis between survey information and observed performed during a simulated small bowel repair shows that intraoperative skills had significant correlation to their perceptions of skill decay and task difficulty during the bowel repair. The findings suggest that evaluating resident’s directional instructions may provide an additional individualized intraoperative assessment metric.

Work in Progress:

The greatest difficulty in evaluating performance of the small bowel repair station is that there are multiple ways to repair an injured portion of the bowel. In a literature search, we were able to identify at least five methods of repair. In order to further understand what methods are acceptable in our context, a survey was developed, reflective of technical and decision skills, and distributed to board-certified surgeons specializing in abdominal surgery. The survey results were then coded and evaluated for consistency. Items resulting in inconsistency of rating resulted in the item being excluded from the grading process. Participant data was then graded by the developed survey. We are now in the final steps of this analysis, attempting to validate our results. Theoretically if the validation works, this would be a valid method to measure decay and progression of surgeons (*Objective Four*).

Urinary Catheterization

Figure 13: Summer 2014 pilot participant of urinary catheterization scenario. General feedback we received from participants was that the simulation was not realistic to what a real life situation would present.

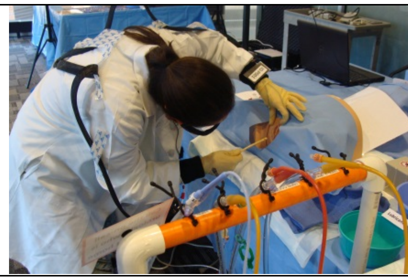


Figure 14: Summer 2015 participant of urinary catheterization scenario. General feedback we received from participants in 2015 was that the simulation was still not realistic to what a real life situation would present.

Figure 15: Summer 2015 setup of urinary catheterization scenario. We required our participants to unscrew their catheter choice from the stand; this caused several researcher errors which were correct in late 2015 into 2016.

Figure 16: Summer 2016 setup of urinary catheterization scenario. Additional cameras were added to best capture catheter insertion.

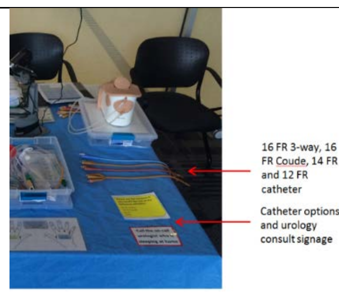


Figure 17: Summer 2016 setup of urinary catheterization scenario. Catheter options were made available for participants to select and a research assistance would hand the entire catheter kit to the participant.

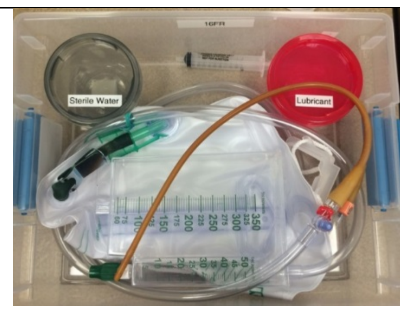


Figure 18: Summer 2016 catheter kit. Participants were given this kit once they selected the catheter size. This best represented the protocol found in a procedure room.

Four simulated scenarios were created in 2014 for patients requiring a urinary catheter. Residents performing this simulation in 2014 were presented with three of the four simulated scenarios: a 27-year-old female presenting as a Level I trauma following a motor vehicle collision, a 45-year-old pre-operative female with a complete bowel obstruction, a 75-year-old male with rectal cancer who is in the operating room for a planned low anterior resection, a 67-year-old post-operative male with benign prostatic hyperplasia following an open cholecystectomy. Five Foley catheter types and sizes were available to the participants, as well as all other necessary instruments. In 2015, simulated scenarios were limited to a 45-year-old pre-operative female with a complete bowel obstruction, and a male with potential pelvic trauma. This allowed for a more participation on each scenario resulting in higher power. In addition, instruments were re-

organized into a catheter kit to better simulate a real-life scenario. These alterations are further detailed in the study's 2015 annual report. In 2016, participants returned to the four simulated scenarios from 2014 to gain a more diverse and comprehensive dataset on the different patient cases. During urinary catheter insertion in all three years of the study, motion metric data was collected from the participant's hands, as well as video and audio recordings. A researcher recorded technical errors made by the participant during task performance for later review. After urinary catheter insertion on the simulation models, participants were asked to describe potential difficulties and describe catheter placement in two of four clinical scenarios. Participant responses were audio recorded and later transcribed for decision-making analyses. Finally, the participant completed a brief survey where they were asked to indicate their confidence in their performance, as well as their perceived difficulty of the task.

The following are publications containing analyses from this simulation followed by a brief description of the discovery made in the publication:

- O'Connell-Long, B., Ray, R., Pugh, C. (2015). Unexpected errors and decreased confidence during bladder catheterization: Are residents ready for complex scenarios? Presented at the Wisconsin Surgical Society Conference, Kohler, WI.

  Results were again presented during a Plenary Session at the 2016 American Surgical Society meeting and later published in the Journal of Surgical Research (see reference below).

- O'Connell-Long, B. R., Ray, R. D., Nathwani, J. N., Fiers, R. M., & Pugh, C. M. (2016). Errors in bladder catheterization: Are residents ready for complex scenarios? *Journal of Surgical Research, 206*(1), 27-31.

  This study investigated whether surgical residents have successfully mastered bladder catheterization skills. We hypothesized that surgical residents would be overly confident in their abilities and underestimate the potential for case complexity. Average pre-simulation confidence was 4.42 (range 1 to 5, standard deviation [SD] = 0.85). 36% of the residents ranked their pre-simulation confidence in problem solving abilities as "moderately confident" or below, whereas 64% were "very confident" or "extremely confident". The lower the resident's pre-simulation confidence in problem-solving, the more errors they committed during the simulation (beta=-0.33, t=-2.15, P=0.04). Participants with higher pre-simulation confidence in problem-solving abilities took less time before deciding to place a urology consult, as well (beta=-1.53, t=-4.32, P=0.001). Average post-simulation confidence was 3.56 (SD=-.81) indicating that the simulation was more difficult than anticipated. Participants made an average of 5.1 errors (SD=2.6). The results of this study demonstrate a need to develop a culture in surgery where residents are able to address gaps in knowledge and practice basic procedures until they reach a level of mastery. Simulation is a valuable tool that allows us to identify areas for improvement and expose potential errors that may occur during complex presentations of basic procedures.

- Nathwani, J. N., Law, K. E., Ray, R. D., O'Connell-Long, B. R., Fiers, R. M., D'Angelo, A. L., DiMarco, S. M., & Pugh, C. M. (2016). Resident performance in complex simulated urinary catheter scenarios. *Journal of Surgical Research, 205*(1), 121-126.

  The aim of this study was to assess the surgical trainee's ability to insert and troubleshoot difficult urinary catheterization scenarios in the setting of common and complex urinary pathology. It was hypothesized that during urinary catheterization, residents would make inconsistent decisions relating to catheter choices and clinical presentations. It was found that for all scenarios where bladder catheterization was achievable, participants were most likely to achieve successful bladder catheterization on their first attempt with their first urinary catheter choice. For the preoperative female, over 60% of participants successfully catheterized the model on their first attempt with their first catheter choice ($X^2_{(2)} = 43.03$, $P<0.0001$). For participants who failed and made a second attempt with the same catheter, the success was approximately 10%. Even fewer participants had success when switching to the second catheter choice at 5% or less. Despite the expectations of the ACS/APDS Surgical Skills Curriculum, it appears that our residents struggle in performance of urinary catheterizations and decision-making. Our study suggests that surgical residents have limited understanding on how to troubleshoot a failed urinary catheterization. Use of simulation and development of standardized algorithms to help develop and maintain resident skill could provide guidance in straightforward scenarios and increase the chances of success.

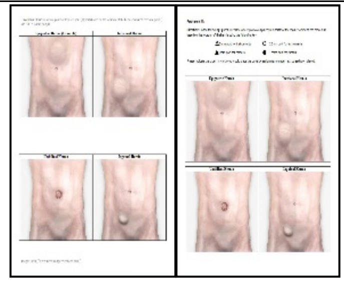
- Nathwani, J. N., Law, K. E., Witt, A. K., Ray, R. D., DiMarco, S. M., & Pugh, C. M. (2016). A simulation-based, cognitive assessment of resident decision making during complex urinary catheterization scenarios. *The American Journal of Surgery.* (Accepted).

  This study explores general surgery residents' decision-making skills during complex urinary catheter scenarios. It was hypothesized that residents would make inconsistent decisions for clinical scenarios where they suspected pathological conditions. In a male trauma scenario, analysis revealed that 55% (N=22) chose to proceed with catheterization, whereas 45% of participants (N=18) chose to place a Urology consult immediately ($\chi^2=0.40$, $p=0.527$). For those who chose to catheterize, 47.5% (N=19) vocalized the decision to gather more information by running a diagnostic test. The majority of those who chose to run diagnostic tests vocalized the need for a retrograde urethrogram (32.5%, N=13, $\chi^2=36.55$, $p<0.001$). Only 7.5% (N=3) of participants chose to place a catheter without first ordering an ancillary test. In the geriatric male scenario, there was no consensus amongst the residents at the group or PGY level after the first decision. In the case of initial catheter failure, participants randomly chose to upsize or downsize catheters or to consult Urology. This study emphasizes two important points. First, simulation provides optimal means to evaluate clinical decision making. Second, residents appear to have little experience and training in catheterization scenarios for which they may be responsible. This information enables opportunities for change in surgical education that would emphasize important clinical topics, continual assessment, and methods for feedback.

Work in Progress:

Due to the changes in simulation stations used in 2015 and 2016, it is difficult to compare resident performance data longitudinally. Because many of the assessed performance errors remained the same no matter which simulation model was used, we aim to compare performance based on common error data (*Objective Two*). In addition, resident's motion data for catheter insertion can be assessed in comparison to expert data to illuminate any discrepancies between resident and expert performance. Finally, decision-making abilities previously analyzed from the cognitive scenario responses in 2014 can be expanded upon (*Objective One*). Longitudinal decision-making abilities will be analyzed using the decision-making tree created by Jay Nathwani in 2016 to elucidate any trends during resident's dedicated research fellowships (*Objective Two*).

Laparoscopic Ventral Hernia Repair

		
Figure 19: Setup of the laparoscopic ventral hernia repair. This simulation had been previously validated so little updates were necessary between years 2014 – 2016.	*Figure 20: Participant completing the laparoscopic ventral hernia repair in summer, 2015.*	*Figure 21: Example of 2014 and updated version of the hernia port placement survey.*

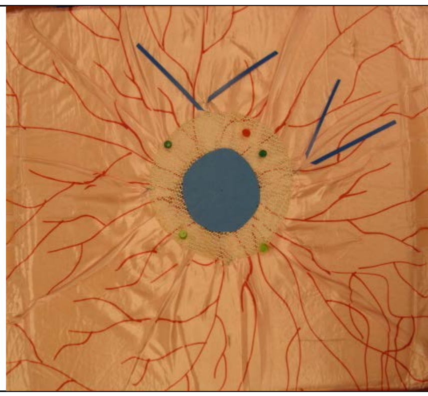
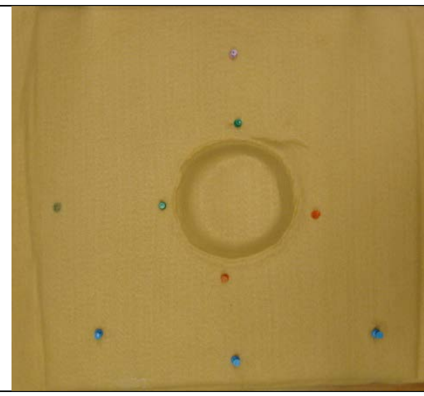
	

Figure 22 and 23: Example of completed final product skin from laparoscopic ventral hernia station

The laparoscopic ventral hernia repair was also designed with the intent of being a challenging task for our participants. Before participating in the simulated scenario, residents were asked to complete a hernia port placement survey. This survey was randomized so that participants were presented with three of four clinical scenarios for an epigastric hernia, an incisional hernia, an umbilical hernia, and an inguinal hernia. In 2014, participants were asked to mark on the photos of the different hernias where they would place ports. In 2015 and 2016, the survey was altered so that participants were required to indicate not only the location of the ports they would place, but also the size and type of the port, and the order in which they would place the ports. Following the survey, the residents were able to begin the simulated scenario. Participants were presented with the following clinical vignette: A 69-year-old male, status post exploratory laparotomy and bowel resection one year ago. The participant is now operating on the patient for an incisional hernia repair. When participants are presented to the scenario, the case has already started and two of four anchoring sutures are brought up through the abdominal wall. Participants are also told that the hernia is 10 x 10 cm, the abdomen is insufflated, ports are in place, and the provided mesh has a 16 cm diameter. Participants are encouraged to complete the procedure by bringing up the remaining two anchoring sutures and placing 5 tacks. The

participants are provided with laparoscopic instruments and an operative assistant at the level of the medical student. Data collection occurred in the form of checklist, video, discourse, motion, and final product. Once collected, checklist and video data were coded for technical and cognitive performance, discourse data was transcribed and final products were graded. Post-performance, participants immediately filled out a survey reflective of confidence with performance with the procedure. Data was collected in the years 2014, 2015, 2016.

The following are abstracts and publications containing analyses from the data collected followed by a brief description of the discovery made in the publication:

- Law K. E., Ray R. D., & Pugh C. M. (2015) A tale of two measures and their association with task outcomes: the case of psychomotor skills and self-assessment. Presented at 2015 AAMC.

  This study sought to identify a relationship between residents' perception of abilities and performance in a simulated procedure. We hypothesized that resident psychomotor performance and their subjective measure of difficulty would significantly correlate to their overall outcome. Surgical residents completed two steps of a simulated hernia procedure and assessed their confidence and perceived difficulty. Hernia skins were evaluated. Psychomotor performance during one procedure step and residents' perceived procedure difficulty significantly correlated to hernia skin quality. A total of forty-six residents participated in the study, with thirty-eight completed the procedure. Residents that placed anchoring sutures correctly (n=25) had significantly higher total scores (r(44)=.315, p=.033). Resident perception of procedure difficulty post repair negatively correlated to their overall completion score (r(44)=-.420, p=.002) meaning that the higher the perceived difficulty, the worse the residents performed. Resident confidence in completing the procedure post repair also positively correlated to their completion score (r(44)=.272, p=.034). From this data, we concluded that psychomotor performance on one procedure step may predict overall outcomes.

- Law, K. E., Jenewein, C. G., Gannon, S. J., DiMarco, S. M., Maulson, L. J., Laufer, S., & Pugh, C. M. (2016). Exploring Hand Coordination as a Measure of Surgical Skill. *Journal of Surgical Research, 205*(1), 192-197.

  The study aim was to identify residents' coordination between dominant and non-dominant hands while grasping for sutures in a laparoscopic ventral hernia repair procedure simulation. We hypothesize residents will rely on their dominant and non-dominant hands unequally while grasping for suture. Surgical residents were video recorded for manual coordination events during the active suture grasping phase of a laparoscopic hernia repair. Manual coordination events were video recorded and then coded by active motion of dominant, non-dominant, or both hands; and manual or unimanual manipulation of hands. Data were evaluated from a total of thirty-six residents. Bimanual coordination was used most (40%) and required the most time on average (M = 20.6 s, SD = 27.2), while unimanual non-

dominant coordination was used least (2.2%; M = 7.9 s, SD = 6.9). Residents relied on their dominant and non-dominant hands unequally ($P < 0.001$). During 24% of events, residents depended on their non-dominant hand ($n = 120$), which was predominantly used to operate the suture passer device. We concluded that residents actively coordinate between dominant and non-dominant hands in nearly half of the attempts of complete suture grasping. Bimanual grasping tasks took longer than other tasks on average, suggesting these tasks were characteristically longer, or switching hands required a greater degree of coordination. This publication motivates future work to occur exploring how task completion time and overall performance are affected by hand utilization.

- Mohamadipanah, H., Parthiban, C., Maulson, L., Laufer, S., Rutherford, D., Law, K., DiMarco, S., & Pugh, C. (2016). Is indecisiveness linked to hand movements in performing laparoscopic ventral hernia repair? *1st International Workshop on Surgical Data Science.* Heidelberg, Germany.

  We aimed to investigate whether indecisiveness in VR is linked to the hand movements during a LVH repair procedure. Data from general surgery residents (N=34) were analyzed. It was found that long periods of indecisiveness in discriminating forces within the VR environment showed a significant positive correlation ($p < 0.05$) with longer translational and angular path lengths of the participants' dominant hand while performing LVH repair. Long periods of indecisiveness were also positively correlated with greater working volumes when repairing a ventral hernia ($p < 0.05$). These findings indicate Path Length and Working Volume as two hand motion metrics that can potentially play a role in the objective assessment of the psychomotor skills necessary to successfully complete a LVH repair procedure.

- Mohamadipanah, H., Parthiban, C., Law, K., Nathwani, J., Maulson, L., DiMarco, S., & Pugh, C. (2016). Hand smoothness in laparoscopic surgery correlates to psychomotor skills in virtual reality. *2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN),* 242-246.

  The main purpose of this study was to find possible relationships between the smoothness of hand function during laparoscopic ventral hernia (LVH) repair and psychomotor skills in a defined virtual reality (VR) environment. Thirty-four surgical residents N = 34 performed two scenarios. First, participants were asked to perform a simulated LVH repair during which their hand movement was tracked using electromagnetic sensors. Subsequently, the smoothness of hand function was calculated for each participant's dominant and non-dominate hand. Then participants performed two modules in a defined VR environment, module 2 and module 4. More smooth hand function during the LVH repair correlated positively with higher performance in VR modules. Also, translational smoothness of dominant hand is found as the most informative smoothness metric in the LVH repair scenario. Therefore, defined force matching and target tracking assessments in VR can potentially be used as an indirect assessment of fine motor skills in the LVH repair.

- Law, K. E., Nathwani, J. N., DiMarco, S. M., Ray, R. D., Gannon, S. J., D'Angelo, A. D., Wiegmann, D. A., & Pugh, C. M. (2016). An Exploration of Moonlighting Effects on Surgical Skill in Lab Residents. Presented abstract at as a quickshot at The Academic Surgical Congress 2016. Manuscript will be submitted to the American Journal of Surgery.

  This study aimed to objectively measure the impact of moonlighting shifts on research residents' operative performance. We hypothesized that moonlighting frequency will positively correlate with laparoscopic skill performance regardless of shift specialty. Forty research residents were given 15 minutes to complete a brief laparoscopic ventral hernia repair. Monthly moonlighting shifts were defined as low frequency or high frequency. Hernia performance and moonlighting shifts were analyzed with a logistic regression. The logistic regression distinguished between the two groups; residents with high moonlighting shifts and better repair scores compared to those with low moonlighting shifts. The results from this study suggest that moonlighting may provide learning benefits, regardless of the type of moonlighting, by allowing participants to maintain and improve skill during research time.

- Nathwani, J. N., Law, K. E., Wise, B. J., Lian, S., Garren, M. E., DiMarco, S. M., & Pugh C. M. How Does Surgical Performance Change in the Laboratory Years? Accepted as an oral presentation at the 2017 Academic Surgical Congress. 2017 planned manuscript submission to *Surgery*.

  This study examines the impact of the academic development years on residents' operative performance. We hypothesize that performance will improve as laboratory residents progress through their academic years. Surgical performance data were collected from laboratory years in the 2014, 2015, and 2016 academic years. Residents performed an abbreviated laparoscopic ventral hernia repair. Final products were graded and compared from the first and second year of participation. A paired t-test shows significant improvement from first year of participation versus second year of participation (14.04 vs 17.22, p=0.022). The scores representative of improvement of hernia score suggest that laboratory resident's surgical skills improve during times of academic development, despite exposure to limited clinical experiences. We speculate that the performance improvement may be a reflection of personal efforts, program-specific call duties, acquisition of moonlighting opportunities, or artifact of repeated exposure to a simplified surgical task.

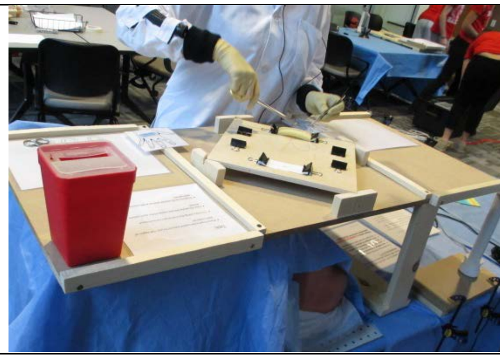
Work in Progress:

We are currently determining how to code and analyze performance on the port placement survey to determine its validity as a training and assessment tool (*Objective Four*). Resident port placement choices for all scenarios will be compared to expert responses to the same scenarios. Longitudinal data is currently being coded for error management in the resident LVH repair performance data. The commitment, detection, and recovery of errors throughout the three years

of the study will be analyzed for trends in clinical performance of the LVH repair (*Objective Three*).

Simple Suture Task

Figure 24: Researcher in summer 2015 reading scenario of simple suture to participant



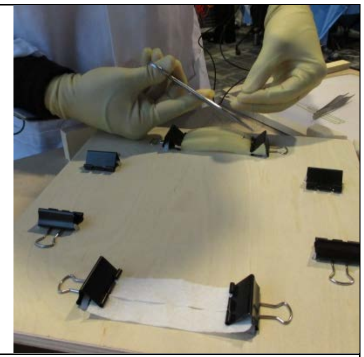


Figure 25 and 26: Participant in 2015 completing the simple suture task on variable simulated tissue. This task allowed researchers to get a baseline understanding of participants basic skills.

A simulated suture task was added in 2015 and 2016 data collections. Participants were presented with two small pieces of foam and two small pieces of paper tissue. They were asked to place three interrupted sutures to join the edges of each material using the instrument tie technique. Four throws/knots were required for each suture placed, however the participants could place the knots wherever they considered appropriate on the materials. The foam and paper tissue segments were used to simulate relatively sturdy and delicate tissue types, respectively. During the simple suture tasks, motion metric data was collected from the participant's hands, as well as video and audio recordings. There was no associated survey data collected on confidence and perceived difficulty for this task.

Work in Progress:

Motion data is being coded into time segments to compare working volume and idle time of resident suturing. In addition, participant's sutured materials are being analyzed for potential final product scores. These scores will be compared to the participant's time away from clinical practice to see if there is any skills decay related to technical skills such as suturing (*Objective Two*). In addition, resident performance will be compared to expert performance in this task (*Objective Three*).

Virtual Reality Skills Assessment Using a Force Dimension Omega Haptic Device

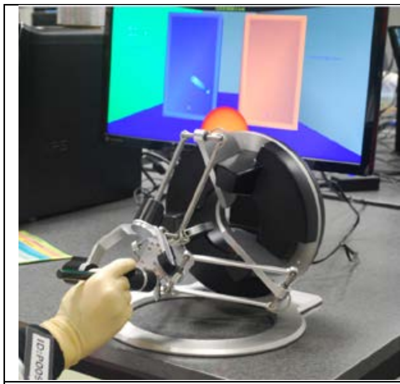


Figure 27: In this figure the participant is using the stylus to tap on the blocks to determine which block is stiffest. Clicking the home button at this time would select block A as the stiffest.

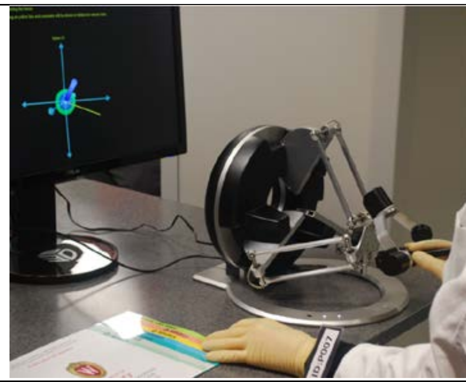


Figure 28: A depiction of the sphere displayed when the stylus is properly centered in the sphere.

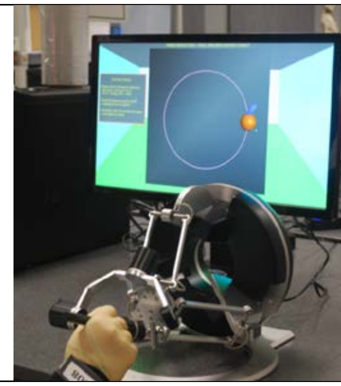


Figure 29: A sample of the target tracking module.

The ability of a surgeon to determine force magnitudes, match opposing forces while in motion or stationary, and reaction time are investigated using a haptic device (Force Dimension Omega ®). The Omega Haptic Robot is a 6 degrees of freedom positioning device with force feedback in 3 degrees of freedom and a passive wrist. The haptic device features a stylus with a home button for the participant to control the virtual environment. Forces are applied to the stylus in various magnitudes and directions, responding to movements and collisions in the VR environment. The four modules designed to assess participant performance using this device are detailed below.

Module 1: Force Discrimination and Just Noticeable Difference Tasks

The goal of this module was to identify a minimum stiffness difference a participant is able to perceive between two given objects. This module was aimed at investigating the correlation between surgical skills and the ability to discriminate forces. The environment was set up to have two blocks of varying stiffness's and the participant guessing the harder material by tapping on each block. In the virtual environment, once a decision is made, the participant directs the stylus to the side of the screen with the chosen block, and clicks on the home button. In some trials, blocks A and B exhibit the same opposing forces, in which case the stylus is directed between the blocks. A total of 19 trials were performed and were set to 15,20, or 30 second time limits. A countdown of time remaining is displayed on the monitor to force the participants to make a decision before the program times out.

Module 2: Stationary Force Matching

In addition to the force discrimination task, a force matching module was developed to study one's ability to resist/match an applied force. The task was separated into two different modules, involving a stationary target and a moving target. The participant must resist opposing forces of varying magnitudes and direction while keeping the stylus centered at a defined, sometimes moving, location. To begin, the participant directs the stylus into the center of the sphere displayed on the monitor. When the stylus is appropriately centered, the sphere will turn green and blue, as shown in Figure 2. The environment will then apply force in a push or pull fashion while the participant attempts to resist these forces. Each participant completes 20 trials. We predicted that one's ability to resist and recover from an applied force would correlate with surgical performance. Based on the motion data acquired from these modules, metrics were developed to compare surgical performance from other clinical simulations in this study.

Module 3: Reaction Time

The goal of this module was to study how fast a person reacts to an applied force as well as the minimum force a person is able to perceive. We wanted to investigate the correlation between reaction time and the minimum force perception to that of the force discrimination task and estimate any decay with time and its corresponding effects on performance. We hypothesized a strong correlation between stiffness discrimination and minimum force perception. The task developed was similar to module 2. A force was applied to the hand of the participant by the haptic device and as soon as the participant realizes a force is applied, they click the home button of the stylus. At that point the time from start to time of the click is recorded as "Reaction Time", and the applied force drops to zero. A total of 19 trials are completed by each participant, with a variety of applied magnitude and directions of the forces.

Module 4: Target Tracking Force Matching

As opposed to a stationary target this module was developed to study the participant's ability to resist force disturbances while objects were in motion. While the stationary target helped to develop metrics, following moving targets were helpful in emulating motion similar to the surgical environment. To keep things simple, we restricted to planar motion and we chose a circular trajectory for the tracking. The goal is to keep the tip of the stylus centered in the sphere moving in a circle as various forces are applied, Figure 3. Each participant completed 15 trials of this task, each with a unique set of magnitudes and directions. The angular velocity of the moving target was varied across the trials and the participant's ability to compensate was recorded.

The following are abstracts and publications containing analyses from the data collected followed by a brief description of the discovery made in the publication:

- Parthiban, C., Ray, R., Rutherford, D., Zinn, M., & Pugh, C. (2016). Development and analysis of psychomotor skills metrics for procedural skills decay. *Medicine Meets Virtual Reality 22, 220*, 285-288.

  We developed and analyzed the metrics associated with a force production task involving a stationary target with the help of advanced VR and the Force Dimension Omega 6 haptic device. We study the effects of force magnitude and direction on the various metrics namely path length, movement smoothness, velocity and acceleration patterns, reaction time and overall error in achieving the target. Data was collected from 47 participants who were residents. Results show a positive correlation between the maximum force applied and the deflection error, velocity while reducing the path length and increasing smoothness with a force of higher magnitude showing the stabilizing characteristics of higher magnitude forces. This approach paves a way to assess and model procedural skills decay.

- Mohamadipanah, H., Parthiban, C., Law, K. E., Dimarco, S., & Pugh, C. M. (2016). Use of simulation as a screening tool for surgical skills: Are we there yet? Presented at: *American College of Surgeons 2016.* Washington, D.C.

  We hypothesized that baseline screening with a Virtual Reality (VR) module will predict final operative quality for a 10 cm ventral hernia repair. PGY 1-5 surgery residents ($n = 46$) were screened using a multi-task VR module prior to completing a simulated Laparoscopic Ventral Hernia (LVH) repair. The VR tasks tested module 2 and module 4. LVH repair quality was assessed using a validated, 10-item checklist. Dependent variables included mesh planning (number of anchoring sutures, quality of suture placement and suture location) and coverage of the hernia defect (percent hernia coverage, distance between end of mesh and fascia and mesh smoothness relevant to the abdominal wall). Despite the PGY range, there were no significant correlations between year in training and LVH completion quality ($p = 0.43$). However, lower scores on the ten item scoring rubric for LVH repair quality, were significantly correlated with poor motor control in module 2 ($r = -0.37$, $p = 0.0114$) and poor recovery from forced errors in module 4 ($r = -0.40$, $p = 0.0058$). VR screening can predict final performance in LVH repair. Decreased costs and significant improvements in VR assessment techniques will eventually allow objective, unbiased screening for LVH repair performance.

Virtual Reality Skills Assessment Using Phantom Omni Haptic Device

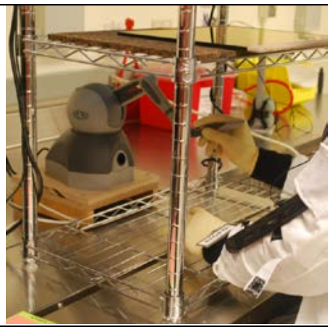
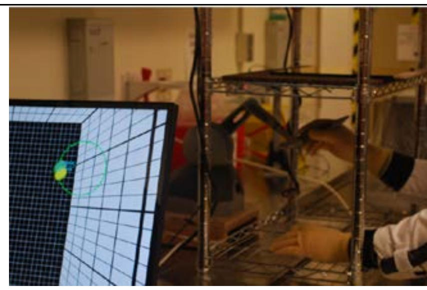
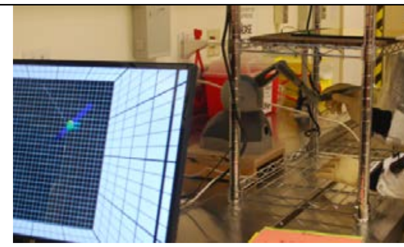
		

Figure 30: The module set-up for station F. Participants handle the stylus from the haptic device below while gazing down into the mirrored images.	*Figure 31: In the foreground an example of the circular path for Module 1 of Station F. In the background the participant can be seen utilizing the stylus of the haptic device to interact with the VR environment.*	*Figure 32: This figure displays the image viewed by the participant in the mirror, as well as another view of the station set-up.*

A second virtual reality assessment was developed for psychomotor skill measurement. The haptic device Phantom Omni, Geomagic Touch ® was selected for these modules. At the station, the participant is seated so they may look down into the mirrored image, displayed in Figure 4. Projecting the display over the workspace where their hands are located creates the sense of working within their line of sight. The objects in the VR environment are matched with the workspace of the stylus. The four modules designed to assess participant performance using this device are detailed below.

Module 1: Training Experience vs. Sensorimotor Task Performance

We investigated to what extent collected survey information about the training level of surgical residents could be recovered from the analysis of performance data collected from each participant (Huang, et al., 2016). This method builds on our previous work of identifying a surgeon's training year by further segmenting participants into key groups: post graduate year (PGY), research years (RY), and clinical years (CY). We devised an interactive virtual environment in which participants could perform multiple testing modules, including movements within visual distortion, force perturbation, and a simulated tissue puncture task. In addition, our experimental apparatus featured physical stations that simulated essential clinical procedures, including laparoscopic ventral hernia (LVH) repair, bowel anastomosis, central venous catheterization, and bladder catheterization. We performed a multivariate analysis of variance (MANOVA) to determine a linear combination of metrics and transformation matrix that best separates the defined population groupings. Next, using data transformed according to MANOVA, we performed a linear discriminant analysis with cross-validation (90% training, 10% testing) to relate the class type to the selected metrics. Our results showed successful classification using combined metrics from the sensorimotor tasks, and separately from the combined metrics from all clinical scenarios. This establishes procedures for identifying the proficiency of surgeons relative to common standards.

Module 2: Movement Identification

Our analysis of identification of surgical training level has demonstrated that metrics on gross movement characteristics from clinical simulators appear to be not as robust as sensorimotor tasks from virtual reality tasks. We have further developed analysis tools to automatically identify segments of goal directed movements during clinical scenario tasks. The goal of this analysis is to refine our metrics of performance in the clinical simulation scenarios so that they reflect the movement most relevant to surgical skill. Our current tool successfully isolates portions of movement that are stereotypically associated with suturing action, then transforms the movement data into a common coordinate frame, and finally applies metrics of smoothness and planar deviation as measures of performance. The contribution of this work to clinical medicine is that it will provide tools to more accurately assess the skills of surgeons particularly in laparoscopic manipulation through the measurement of movement kinematics of the hand alone.

Publications cited above:

- Huang, F. C., Mussa-Ivaldi, F. A., Pugh, C. M., Mohamadipanah, H. (2016). Identification of surgical training level with data from clinical scenario simulation and sensorimotor tasks. *Society for Neuroscience 2016 Conference.* San Diego, CA.

Work in progress:

Cross correlation analysis between virtual reality modules and metrics from medical simulators (Central Line, Bowel Repair, Urinary Catheterization, and LVH repair) are in progress. In addition, analysis of the differences in performance of the participants in consecutive years is also under investigation.

Final Conclusions

The conclusion time period of the project has included a number of significant steps towards meeting the four key study objectives outlined by our original SoW. Specifically, the team has successfully completed the third year of data collection; collected data from retired surgeons; analyzed performance on multiple stations; and disseminated our preliminary work in the form of several papers and presentations. We used the no cost extension to finalize our data analysis plan and move towards completing work on all four areas of our SoW will be imperative to our research. There is still work in progress being completed as noted throughout the report; and we plan to continue analyzing and publishing high impact papers based on our findings.

Appendix of Submitted Abstracts

1. D'Angelo, A. D., Ray, R. D., Jenewein, C. G., Jones, G. F., & Pugh, C. M. (2015). Residents' perception of skill decay during dedicated research time. *American Surgical Society.*
2. Law, K. E., Ray, R., & Pugh, C. M. (2015). A tale of two measures and their association with task outcomes: The case of psychomotor skills and self-assessment. *Association of American Medical Colleges.*
3. Gwillim, E. C., Law, K. E., & Pugh, C. M. (2015). Error tolerance: A new analysis approach in laparoscopic surgical simulation. *Association of American Medical Colleges.*
4. Gwillim, E. C., D'Angelo, A. D., Law, K. E., Cohen, E. R., Rutherford, D. N., & Pugh, C. M. (2015). Error tolerance: A new psychomotor performance metric in laparoscopic surgery. *American College of Surgeons-Accredited Education Institutes.*
5. O'Connell-Long, B., Ray, R., & Pugh, C. (2015). Unexpected errors and decreased confidence during bladder catheterization: Are residents ready for complex scenarios? *Wisconsin Surgical Society.*
6. Gannon, S., Law, K., Ray, R., D'Angelo, A., DiMarco, S., & Pugh, C. (2016). Do resident's operative leadership skills correlate with self-assessments of technical skill? *Academic Surgical Congress.*
7. O'Connell-Long, B., Ray, R., Nathwani, J., Fiers, R., & Pugh, C. (2016). The relationship between confidence level and procedural errors: A bladder catheterization model. *Academic Surgical Congress.*
8. Nathwani, J. N., Law, K. E., Ray, R. D., O'Connell-Long, B. R., Fiers, R. M., D'Angelo, A. D., DiMarco, S. M., & Pugh, C. M. (2016). Resident performance in complex simulated urinary catheter scenarios. *Academic Surgical Congress.*
9. Law, K. E., Rutherford, D. N., Gannon, S. J., Millar, C. T., & Pugh, C. M. (2016). Hand dominance and coordination during suture grasping in a simulated laparoscopic hernia repair. *Academic Surgical Congress.*
10. Law, K. E., Gannon, S., D'Angelo, A. D., Wiegmann, D., & Pugh, C. M. (2016). An exploration of moonlighting effects on surgical skill in lab residents. *Academic Surgical Congress.*
11. Millar, C., Rutherford, D., Law, K., Ray, R., & Pugh, C. (2016). How motion tracking related to end product quality in laparoscopic procedures. *Academic Surgical Congress.*
12. Nathwani, J. N., Fiers, R., Ray, R. D., & Pugh, C. M. (2016). The relationship between technical errors and decision making skills in the junior resident. *Association of Program Directors in Surgery.*
13. Law, K. E., D'Angelo, A. D., Cohen, E., Ray, R. D., Linsmeier, E., Wiegmann, D. A., & Pugh, C. M. (2016). Exploring changes in senior residents' intraoperative error management strategies as a measure of learning. *Association of Program Directors in Surgery.*
14. D'Angelo, A. D., Law, K. E., Cohen, E. R., Ray, R. D., Shaffer, D. W., & Pugh, C. M. (2016). Error management: Do residents identify operative errors as reversible? *Association for Surgical Education.*

15. Mohamadipanah, H., Parthiban, C., Law, K. E., DiMarco, S. M., & Pugh, C. M. (2016). Use of simulation as a screening tool for surgical skills: Are we there yet? *American College of Surgeons.*
16. Jones, G. F., Law, K. E., Jenewein, C. G., Ray, R. D., DiMarco, S. M., & Pugh, C. M. (2016). Assessment during clinical years changes resident perception of skills decay. *Society of Black Academic Surgeons.*
17. Nathwani, J. N., Law, K. E., Ray, R. D., Witt, A. K., Pugh, C. M. (2016). Junior lab resident's approach to complex urinary pathology. *Society of Black Academic Surgeons.*
18. Law, K. E., Ray, R. D., Linsmeier, E., D'Angelo, A. D., Cohen, E. R., Wiegmann, D., & Pugh, C. M. (2016). Do errors and critical events relate to hernia repair outcomes? *Society of Black Academic Surgeons.*
19. Huang, F. C., Mussa-Ivaldi, F. A., Pugh, C. M., & Mohamadipanah, H. (2016). Identification of surgical training level with data from clinical scenario simulation and sensorimotor tasks. *Society for Neuroscience.*
20. Nathwani, J. N., Law, K. E., Wise, B. J., Lian, S., Garren, M. E., DiMarco, S. M., & Pugh, C. M. (2017). *Academic Surgical Congress*.
21. Witt, A. K., Nathwani, J. N., DiMarco, S. M., & Pugh, C. M. (2017). Skill loss and maintenance for research residents: A longitudinal study. *Society of Black Academic Surgeons.* Submitted.

Appendix of Publications

1. D'Angelo, A. L., Ray, R. D., Jenewein, C. G., Jones, G. F., & Pugh, C. M. (2015). Residents' perception of skill decay during dedicated research time. *Journal of Surgical Research, 199*(1), 23-31.
2. Mohamadipanah, H., Parthiban, C., Maulson, L., Laufer, S., Rutherford, D., Law, K., DiMarco, S., & Pugh, C. (2016). Is indecisiveness linked to hand movements in performing laparoscopic ventral hernia repair? *1st International Workshop on Surgical Data Science.* Heidelberg, Germany.
3. Mohamadipanah, H., Parthiban, C., Law, K., Nathwani, J., Maulson, L., DiMarco, S., & Pugh, C. (2016). Hand smoothness in laparoscopic surgery correlates to psychomotor skills in virtual reality. *2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 242-246.
4. Nathwani, J. N., Law, K. E., Ray, R. D., O'Connell-Long, B. R., Fiers, R. M., D'Angelo, A. L., DiMarco, S. M., & Pugh, C. M. (2016). Resident performance in complex simulated urinary catheter scenarios. *Journal of Surgical Research, 205*(1), 121-126.
5. Law, K. E., Jenewein, C. G., Gannon, S. J., DiMarco, S. M., Maulson, L. J., Laufer, S., & Pugh, C. M. (2016). Exploring Hand Coordination as a Measure of Surgical Skill. *Journal of Surgical Research, 205*(1), 192-197.
6. Mohamadipanah, H., Parthiban, C., Nathwani, J., Rutherford, D., DiMarco, S., & Pugh, C. (2016). Can a virtual reality assessment of fine motor skill predict successful central line insertion? *The American Journal of Surgery, 212*(4), 573-578.
7. Nathwani, J. N., Fiers, R. M., Ray, R. D., Witt, A. K., Law, K. E., DiMarco, S., & Pugh, C. M. (2016). Relationship between technical errors and decision-making skills in the junior resident. *Journal of Surgical Education, 73*(6), e84-e90.
8. O'Connell-Long, B. R., Ray, R. D., Nathwani, J. N., Fiers, R. M., & Pugh, C. M. (2016). Errors in bladder catheterization: Are residents ready for complex scenarios? *Journal of Surgical Research, 206*(1), 27-31.
9. Gannon, S. J., Law, K. E., Ray, R. D., Nathwani, J. N., DiMarco, S. M., D'Angelo, A. L. D., & Pugh, C. M. (2016). Do resident's leadership skills relate to ratings of technical skill? *Journal of Surgical Research*, *206*(2), 466-471.
10. Parthiban, C., Ray, R., Rutherford, D., Zinn, M., & Pugh, C. (2016). Development and analysis of psychomotor skills metrics for procedural skills decay. *Medicine Meets Virtual Reality 22, 220*, 285-288.

Personnel Supported on the Project

PI
Carla Pugh (41 months)

Co-PI's
Lee Faucher (25 months)F
Eugene Foley (30 months)
Caprice Greenberg (32 months)
Jacob Greenberg (10 months)
Douglas Wiegmann (30 months)
Michael Zinn (24 months)

Researchers
Patrick Barlow (6 months)
Eran Gwillim (9 months)
Donna Hankins (23 months)
Caitlin Jenewein (24 months)
Katherine Law (30 months)
Hossein Mohamadipanah (12 months)
Chembian Parthiban (24 months)
Rebecca Ray (20 months)
Andrew Rutherford (22 months)

Administration
Shannon DiMarco (31 months)
Mary Ellen Hagenauer (5 months)
Mara Snyder (3 months)

Hourly Researchers
Danielle Abramov (9 months)
Crystal Chen (3 months)
Lily Comp (3 months)
Michael Daniels (3 months)
Rebekah Fiers (16 months)
Samantha Gannon (21 months)
Anna Garren (6 months)
Maggie Garren (6 months)
Jordan Gates (12 months)
Sining Lian (9 months)
Elyse Linsmeier (12 months)

Development and Analysis of Psychomotor Skills Metrics for Procedural Skills Decay

Chembian PARTHIBAN [a,b,1,2], Rebecca RAY [b], Drew RUTHERFORD [c], Mike ZINN [a], Carla PUGH [b]

[a] *University of Wisconsin-Madison, Department of Mechanical Engineering*

[b] *University of Wisconsin-Madison, Department of General Surgery*

[c] *University of Wisconsin-La Crosse, Department of Health Professions*

Abstract. In this paper we develop and analyze the metrics associated with a force production task involving a stationary target with the help of advanced VR and Force Dimension Omega 6 haptic device. We study the effects of force magnitude and direction on the various metrics namely path length, movement smoothness, velocity and acceleration patterns, reaction time and overall error in achieving the target. Data was collected from 47 participants who were residents. Results show a positive correlation between the maximum force applied and the deflection error, velocity while reducing the path length and increasing smoothness with a force of higher magnitude showing the stabilizing characteristics of higher magnitude forces. This approach paves a way to assess and model procedural skills decay.

Keywords. Virtual reality, force production, force-motion, haptics

1. Introduction

Medical and surgical procedures have a complex, heterogeneous mix of procedural, psycho-motor and cognitive components. When modeling skills decay, it is critical to identify the components that are most vulnerable and the ones that are relatively stable over time. In the past research has shown that discrete, highly-ordered, procedural skills decay more rapidly than psycho-motor or cognitive skills [1-3]. In order to understand the relationship between these components and to assess the individual differences in psycho-motor ability we have developed motor control and perceptual tasks with the help of VR and haptics systems.

This paper focuses on the development and analyses of the varies metrics associated with motor control based on a force production task which can then be used to correlate with cognitive skills analysis developed using error-enable boxed trainer tasks namely, Laparoscopic Ventral Hernia Repair and Central line placement. Using advanced VR and haptic systems we are able to break down such complex procedures into a set of simplified motor control tasks. In this case, the goal of the module was to reach and stay at the center of a stationary target compensating for the external force disturbances which can be used to study the ability to resist and react disturbances. The various metrics that

[1] Corresponding Author: Chembian Parthiban, Department of Mechanical Engineering, University of Wisconsin-Madison, Madison, WI 53706 USA (e-mail: parthiban@wisc.edu)

2 This work was supported by the Department of Defense W81XWH-13-1-0080 Grant titled Psycho-Motor and Error Enabled Simulations: Modeling Vulnerable Skills in the Pre-Mastery Phase.

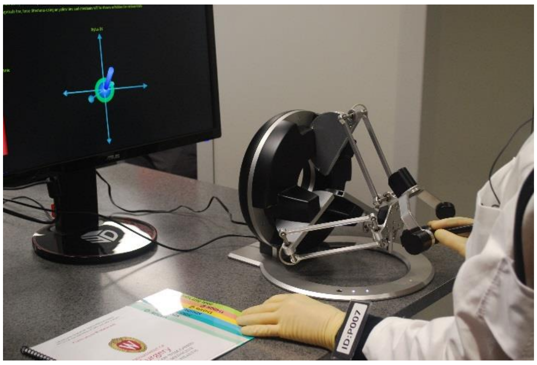
were taken into account in this study were path length, movement smoothness, velocity and acceleration patterns, reaction time and overall error in achieving the target. This provides a strong platform to assess and understand the psycho-motor skills and its contribution while modeling skills decay.

2. Materials and Methods

2.1. Setup and Participants

Fig. 1 shows the VR station setup at a data collection center. The Force Dimension Omega 6 haptic device was used by the participants to perform the virtual reality task. The rendering was created with the help of open source CHAI3D C++ library that supports the Omega device. The device properties are summarized in Fig. 1.

Data was collected from 47 participants during the summer and fall of 2014. Study participants were general surgery residents (post graduate year (PGY) 2-4) engaged in dedicated laboratory time from multiple Midwestern general surgery training programs enrolled in a longitudinal simulation-based assessment study. Data collection occurred at five sites in three Midwestern cities: Madison, WI, Chicago, IL, and Rochester, MN. This study was approved by the University of Wisconsin Institutional Review Board and written informed consent was obtained from all participants.



Workspace	**Translation**	**∅ 160 x 110 mm**
	Rotation	**240 x 140 x 320 deg**
Forces	**Translational**	**12.0 N**
Resolution	**Translation**	**< 0.01 mm**
	Rotation	**0.09 deg**
Stiffness	**Closed-Loop**	**14.5 N/mm**

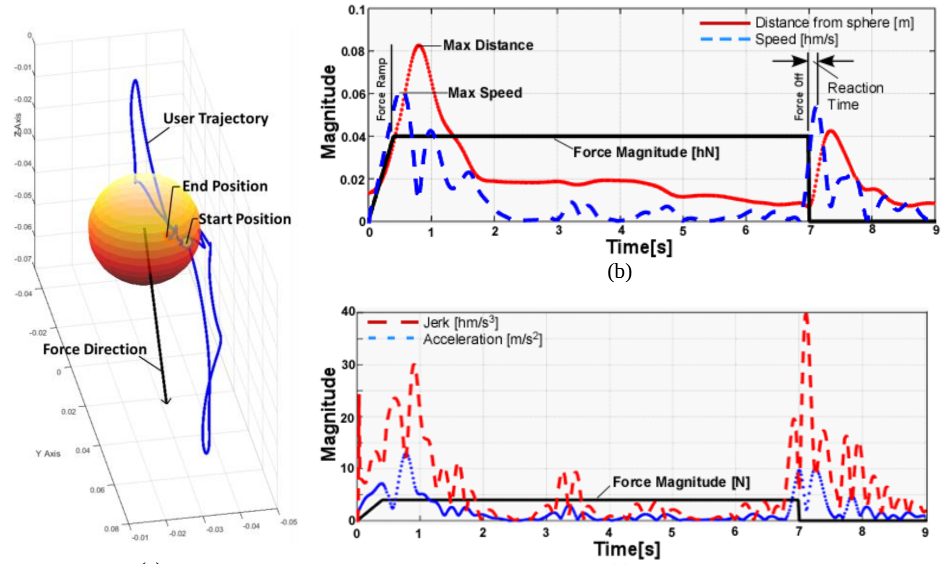
Figure 1. Shows the setup of the VR station during a data collection along with the six degree of freedom Omega device and its parameters.

2.2. Force Production Task

The force production module was developed to analyze the individual difference in psycho-motor abilities specifically with respect to force production and matching capabilities. The task involves placing the tip of the virtual stylus inside a sphere at the center of the world. The goal was to keep the stylus as close as possible to the center of the sphere while forces are being applied. A visual indication is provided by changing the color of the sphere from red to blue when the stylus is tip is inside. Each participant was provided with 20 trials each lasting for a period of 9 seconds. The applied force magnitude was varied between 2 N and 6.5 N and the direction of force was also varied as a push or a pull force from 6 possible different directions. The participant’s global

position, orientation and the forces were recorded for each trial at a sample rate of 200 Hz. A second order, dual-pass Butterworth filter with a cut-off frequency of 7 Hz was used to filter out the high frequency components and the velocity, acceleration and jerk were calculated.

Fig. 2 shows the trajectory and time domain plots for a single participant during a single trial condition. In this case the applied force was pull force of magnitude 4N. Fig. 2(b) shows the trajectory of the stylus tip during the entire trial duration with respect to the sphere and the applied force direction. Fig. 2(a) shows the overlaid time domain plot of the applied force (solid black), the distance from center of sphere (dotted red) and the tip speed (dashed blue). We can observe the peaks during the force application and force release events. Fig. 2(c) shows the magnitude of tip acceleration (dashed blue) and the jerk (solid red) during the same condition. Note that the values showed in the plot are scaled down to show the correlation between the events.



(a) (c)

Figure 2. Shows the time domain plots during a single trial for a single participant. (a) 3D trajectory of the stylus position , (b) Overlaid plot of the distance from sphere, speed and force magnitude and (c) Overlaid plot of the magnitude of the acceleration, jerk and the applied force. (Note: Units adjusted to fit plot)

3. Results

A generalized linear model of the log transformed data demonstrated that increases in the maximum force applied during the trial was associated with several metrics of interest, namely increases log peak deflection force release and increases log peak velocity force release $\beta = 0.24,\ \chi^2 = 576.23,\ p < .001, and\ \beta = 0.26,\ \chi^2 = 599.43,\ p < .001$, respectively). Furthermore, while the force was constantly applied, greater force was associated with shorter log path length and greater log sum jerk $(\beta = -0.069,\ \chi^2 = 33.22,\ p < .001, and\ \beta = 0.06,\ \chi^2 = 42.68,\ p < .001)$. Fig. 3 shows the relationships between the variables with respect to the applied force

magnitude at the time of force release (a) and during constant force application (b). The force release event highlights the participant's ability to react to the sudden change in the applied force and the results are as expected. It's interesting to note that the higher force magnitude stabilizes the movement during the constant force application. Also the reaction time (force release) shown in Fig. 2(a) was not associated with force maximum.

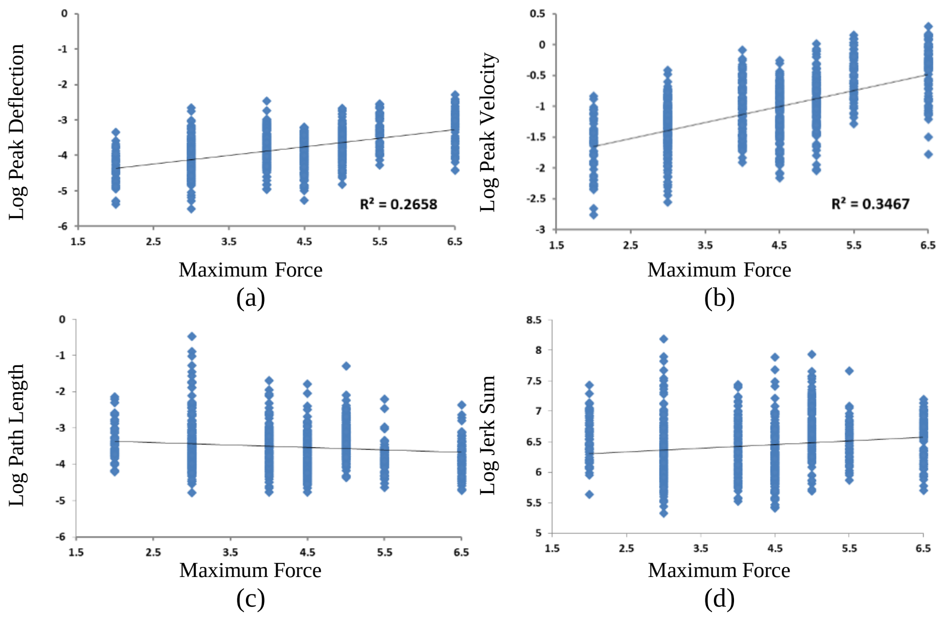


Figure 3. Shows the variation of the log transformed metrics with respect to the maximum applied force. While (a) and (b) correspond to the force release, (c) and (d) are measured during constant force application.

4. Conclusions and Future Work

In this paper the various metrics associated with a force production task involving a stationary target was developed and studied with the help of VR and Force Dimension haptic device. These metrics will be used to correlate with the analyses obtained from three other tasks involving a moving target and just noticeable difference tests to help better understand force perception and its correlation with surgical procedures.

References

Westwood, J. D., et al. "Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST VR laparoscopic simulator." *Medicine Meets Virtual Reality:art,science,technology:healthcare and evolution*(1998): 124.

Oostwoud Wijdenes L, Brenner E, Smeets JBJ. Fast and fine-tuned corrections when the target of a hand movement is displaced. *Experimental Brain Research Experimentelle Hirnforschung Experimentation Cerebrale*. 2011;214(3):453-462. doi:10.1007/s00221-011-2843-4.

Pugh, Carla M., et al. "The Effect of Simulator Use on Learning." *Medicine Meets Virtual Reality 2001: Outer Space, Inner Space, Virtual Space* 81 (2001): 396.

Hand Smoothness in Laparoscopic Surgery Correlates to Psychomotor Skills in Virtual Reality

H. Mohamadipanah[1], C. Parthiban[2], K. Law[3], J. Nathwani[1], L. Maulson[1], S. DiMarco[1], and C. Pugh[1,3]

Abstract—**The main purpose of this study is to find possible relationships between the smoothness of hand function during laparoscopic ventral hernia (LVH) repair and psychomotor skills in a defined virtual reality (VR) environment. Thirty four surgical residents N = 34 performed two scenarios. First, participants were asked to perform a simulated LVH repair during which their hand movement was tracked using electromagnetic sensors. Subsequently, the smoothness of hand function was calculated for each participant's dominant and non-dominate hand. Then participants performed two modules in a defined VR environment, which assessed their force matching and target tracking capabilities. More smooth hand function during the LVH repair correlated positively with higher performance in VR modules. Also, translational smoothness of dominant hand is found as the most informative smoothness metric in the LVH repair scenario. Therefore, defined force matching and target tracking assessments in VR can potentially be used as an indirect assessment of fine motor skills in the LVH repair.**

I. Introduction

Laparoscopic Ventral Hernia (LVH) repair is a frequently performed surgical procedure [1]. A ventral hernia can develop in up to 20% of laparoscopic incisions [2], 24% of open surgery repairs [3], and has an even higher recurrence rate [4]. Given the frequency of occurrence, a high quality education is necessary for proper hernia repair. Present evaluation methods for this procedure are often performed observationally by faculty members through task-specific checklists and global rating scales [5]. However, such methods have shown inconsistent validity evidence and may be prone to bias or subjectivity [6-8].

To overcome this challenge, developing objective assessment tools is imperative. An objective non-human LVH repair assessment can potentially be free of bias and more effective and reliable than any human-based assessment.

An important aspect of surgical skill assessment is the evaluation of psychomotor skills. This is mainly tackled in the literature through defining motion metrics [9], such as path length, working volume [10], and idle time [11]. One example of motion metrics was defined to consider the smoothness of hand function [12-14] during procedures. For instance, in [15], the expert participants had smoother instrument motion than novices when performing a complex suturing task in a laparoscopic simulator. Also, [16] the smoothness metric indicated significant differences between medical students, junior trainees, senior trainees, and consultant surgeons for the three tasks: laparoscopic orientation, object positioning, and sharp dissection.

Although smoothness measures have been used to assess surgical skill, further investigational studies are desired to completely understand the role and usage of psychomotor metrics to assess performance during specified medical procedures. In this study we focused on the LVH repair procedure using a previously validated LVH simulator [17-19]. It has been already shown that the smoothness of hand function significantly correlates to the final quality of the product in LVH repair [20]. Since VR has the capability to assess psychomotor skills, we developed a VR environment to identify the relationship between hand function smoothness in LVH repair and psychomotor abilities in VR. This work also investigates angular smoothness metric, which is defined based on Euler angles, beside the translational smoothness metric in the literature.

In Section II, the description of the method is presented, including the LVH simulation scenario explanation and the description of the VR modules. Then experimental results are presented in Section III and the conclusion and future works are presented in Section IV.

II. Method

A. Setting

Thirty-four general surgery residents from multiple institutions across the Midwest (female = 16, male = 18; Post-Graduate Year (PGY) 1 = 3, PGY2 = 12, PGY3 = 11, PGY4 = 6, and PGY5 = 2) participated in this study. Data collection took place at the designated institution of each study participant. When asked to identify handedness, 31 participants stated they were right-handed, 2 stated they were left-handed, and 1 participant identified as ambidextrous (for the purposes of this study, this participant is assumed to be right-handed).

To measure hand movements, three sensors (tracSTAR Model 180) were affixed to each hand on the participants' index finger, thumb, and wrist, as illustrated in Fig. 1. Medical gloves were then placed on each participant's hands, over the sensors to keep them in place. Sampled at 180 Hz, three translational ($r = [x, y, z]^T$) and three angular variables (Euler angles: $\Theta = [\alpha, \beta, \gamma]^T$) were provided by the tracking system (Innovative Sport Training, Inc., Chicago, IL). A low-pass Butterworth filter (second-order) with a cutoff frequency of 1 Hz [21] was used to remove noise from the data.

[1]H. Mohamadipanah, J. Nathwani, L. Maulson, S. DiMarco, and C. Pugh are with the Department of Surgery, University of Wisconsin-Madison, WI 53726 USA (corresponding author phone: 608-262-2147; e-mail: pugh@surgery.wisc.edu).

[2]C. Parthiban is with the Department of Mechanical Engineering, University of Wisconsin-Madison, WI 53706 USA.

[3]K. Law and C. Pugh are with the Department of Industrial and Systems Engineering, University of Wisconsin-Madison, WI 53706 USA.

978-1-5090-3087-3/16/$31.00 ©2016 IEEE

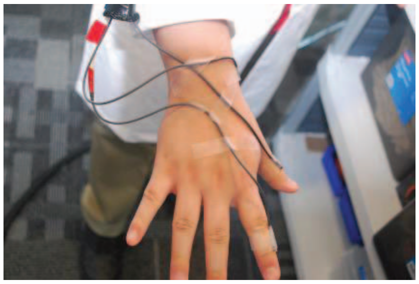


Figure 1. Three sensors were attached to the participant's hand.

"Translational Smoothness", (1), is a metric based on the derivative of translational acceleration [22]. It can be calculated from the following equation, where $\bar{\dot{r}}$ represents the average translational velocity.

$$TS = -\frac{(t_2 - t_1)^3}{(\bar{\dot{r}})^2}\sum_{i=t_1}^{t_2}(\ddot{r}(i))^2 \quad (1)$$

Likewise, "Angular Smoothness", (2), is a metric based on the derivative of angular acceleration, where $\bar{\dot{\Theta}}$ represents the average angular velocity. Higher values of these metrics represent greater smoothness.

$$AS = -\frac{(t_2 - t_1)^3}{(\bar{\dot{\Theta}})^2}\sum_{i=t_1}^{t_2}(\ddot{\Theta}(i))^2 \quad (2)$$

Additionally, psychomotor skills within a defined VR environment were also assessed in this study through the use of the Force Dimension ® haptic device, as shown in Fig. 2. Participants were asked to operate the haptic device using their dominant hands.

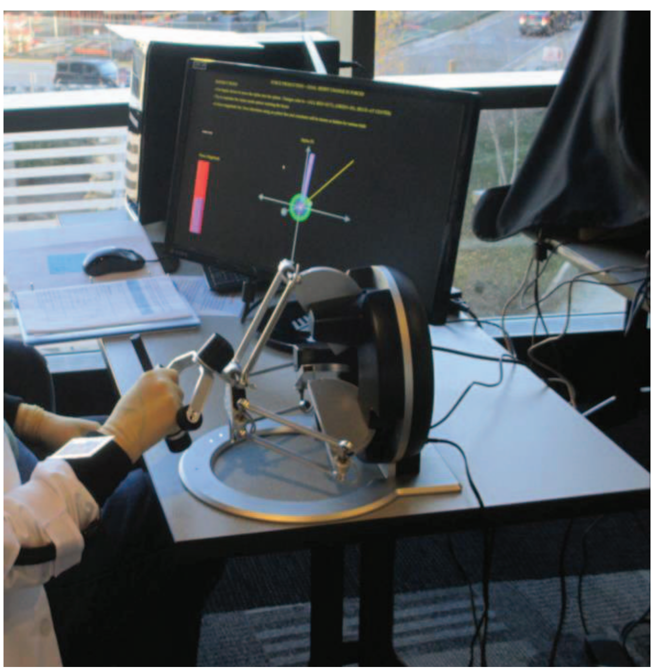


Figure 2. The haptic device being operated by a participant while the participant can see the VR scenario via a monitor.

B. Laparoscopic Ventral Hernia Repair Simulation

Participants were asked to perform a simulated LVH repair, as presented in Fig. 3. The simulator used for this study was previously designed and validated [17-19]. It included an abdominal box with an attached abdominal skin that had a 10cm x 10cm ventral hernia. Except for cautery, all laparoscopic instruments necessary to complete a full LVH repair were provided. Prior to the start of the procedure, researchers pre-placed instrument ports and pulled up two of the four sets of transfascial sutures. Each participant was told that they could assume the abdomen was already been insufflated. Participants were instructed to pull up the remaining two transfascial suture sets within the given 15 minute time limit. Only the first suture set was considered in the motion metric calculations.

Figure 3. LVH repair simulator; a participant performing the procedure while watching the endoscopic camera on a screen.

C. Force Matching in Virtual Reality

Two modules within the VR environment were designed to analyze psychomotor skills. The first module was developed to better understand individual psychomotor differences with respect to force production. More specifically, it aims to investigate an individual's ability to match force and keep a virtual stylus centered upon a stationary sphere while varying forces of different directions and magnitudes are applied. A snapshot of this task is shown in Fig. 4, which includes an inner orange sphere and outer red sphere indicating the participant is outside the target area (the outer sphere turns green and inner sphere turns blue when the stylus is inside the target area, Fig. 2). Participants were asked to complete 20 trials of this task, with force magnitudes ranging from 2 N to 6.5 N and force direction varying in a push or pull fashion. These variations are detailed more specifically for each trial condition in Table III of the Appendix.

Two metrics were defined and investigated within this module. First, "Maximum Distance from Sphere" refers to the 20 trial average of the maximum distance between the tip of the stylus and the center of the sphere. Secondly, "Peak Deflection" refers to the 20 trial average of the maximum stylus deflection from the center of the sphere immediately after the force was applied.

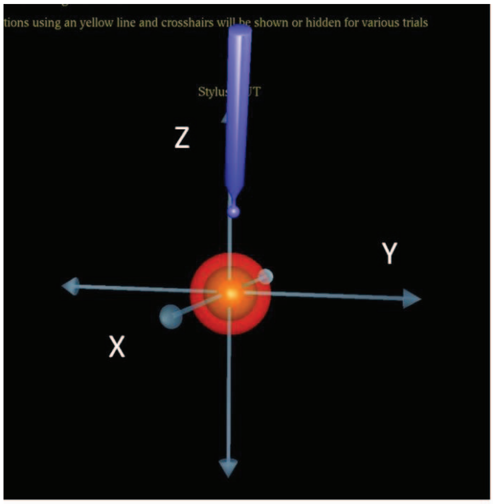


Figure 4. Force matching module in the VR environment. Participants were asked to keep the blue stylus at the center of the orange sphere.

D. Target Tracking in Virtual Reality

The second VR module aims to investigate an individual's ability to track an object while matches forces as well. In this module a virtual stylus centered upon a moving sphere around a circular trajectory with varying forces being applied, Fig. 5. Participants were asked to complete 15 trials of this task, with force magnitudes and directions varying. These variations are detailed more specifically in Table IV of the Appendix.

Psychomotor metrics investigated in this module include: "Path Length", "Maximum Velocity", "Maximum Acceleration", "Maximum Jerk", and "Jerk Summation". All of these metrics were averaged over the 15 trials completed by each participant with "Path Length" being defined as the total travel distance of the stylus tip and "Jerk Summation" being defined as the summation of the jerk for each trial.

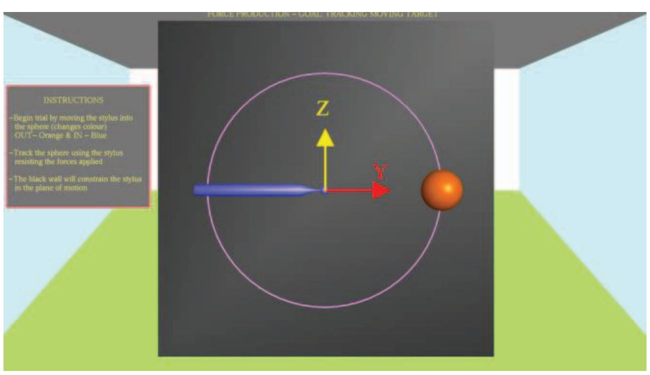


Figure 5. The target tracking module; the participants were supposed to keep the tip of the stylus at the center of the orange sphere while the sphere rotates along the circular path.

III. RESULTS

Video recordings were used to segment out specific durations of the LVH procedure to isolate the securing of the first suture set. The metrics "Translational Smoothness" and "Angular Smoothness" were calculated based only on the securing of the first suture set and were calculated separately for each participant's Dominant hand and Non-Dominant hand.

Spearman's rho correlation calculated between the LVH repair metrics and force matching metrics of the first VR module can be seen in Table I. Participants with more translational smoothness in their dominant hand throughout securing the first suture in the LVH repair procedure showed better performance in the VR force matching module as indicated by less peak deflection and less maximum distance from the sphere. In other words, the LVH metric "Translational Smoothness Dominant Hand" is negatively correlated ($p<0.05$) to the metrics "Peak Deflection" and "Maximum Distance from Sphere" from the force matching module in VR.

TABLE I. CORRELATIONS BETWEEN LVH METRICS AND FORCE MATCHING MODULE METRICS IN VR

Metric of LVH	Metric of Force Matching in VR	R	P value
Translational Smoothness of Dominant Hand	**Peak Deflection**	**-0.3479**	**0.0444**
	Maximum Distance from Sphere	**-0.3653**	**0.0343**
Angular Smoothness of Dominant Hand	Peak Deflection	-0.2587	0.1394
	Maximum Distance from Sphere	-0.277	0.1127
Translational Smoothness of Non-Dominant Hand	Peak Deflection	-0.3008	0.0841
	Maximum Distance from Sphere	-0.3244	0.0618
Angular Smoothness of Non-Dominant Hand	Peak Deflection	-0.2186	0.2133
	Maximum Distance from Sphere	-0.2529	0.1487

Similarly, Spearman's rho correlation calculated between the LVH repair metrics and target tracking metrics in VR can be seen in Table II. All four defined metrics of the LVH repair show significant negative correlations to the VR target tracking metrics. This indicates that smoother hand motions in performing the LVH repair also correlates to increased performance in target tracking. Notably, the metrics "Path Length", "Maximum Jerk", and "Maximum Velocity" all have significant negative correlations ($P<0.05$) with all of the defined LVH smoothness metrics. Furthermore, "Jerk Summation" has significant negative correlations with all of the LVH smoothness metrics except "Angular Smoothness of Non-Dominant Hand". The metric "Maximum Acceleration" only had one significant negative correlation with "Translational Smoothness of Dominant Hand".

In summary, the LVH repair metric "Translational Smoothness of Dominant Hand" was found to be the most informative smoothness metric as this is the only metric that correlated significantly with the force matching metrics. Also, the metrics from the moving target module have shown the lowest p values for the metric "Translational Smoothness of Dominant Hand".

Future work aims to increase the participant sample size, include more residents and expert surgeons, and utilizing more conservative statistical analysis to increase statistical generalizability.

TABLE II. CORRELATIONS BETWEEN LVH METRICS AND MOVING TARGET MODULE METRICS IN VR

Metric of LVH	Metric of Target Tracking in VR	R	P value
Translational Smoothness of Dominant Hand	**Jerk Summation**	**-0.4188**	**0.0143**
	Maximum Jerk	**-0.4824**	**0.0043**
	Maximum Acceleration	**-0.3421**	**0.0483**
	Maximum Velocity	**-0.4662**	**0.0059**
	Path Length	**-0.3812**	**0.0268**
Angular Smoothness of Dominant Hand	**Jerk Summation**	**-0.3595**	**0.0374**
	Maximum Jerk	**-0.3846**	**0.0254**
	Maximum Acceleration	-0.2333	0.1836
	Maximum Velocity	**-0.4124**	**0.0160**
	Path Length	**-0.3439**	**0.0470**
Translational Smoothness of Non-Dominant Hand	**Jerk Summation**	**-0.3586**	**0.0380**
	Maximum Jerk	**-0.4215**	**0.0137**
	Maximum Acceleration	-0.2477	0.1575
	Maximum Velocity	**-0.3528**	**0.0414**
	Path Length	**-0.3675**	**0.0332**
Angular Smoothness of Non-Dominant Hand	Jerk Summation	-0.3109	0.0739
	Maximum Jerk	**-0.4252**	**0.0128**
	Maximum Acceleration	-0.1884	0.2848
	Maximum Velocity	**-0.3409**	**0.0491**
	Path Length	**-0.3653**	**0.0343**

IV. CONCLUSION

The main purpose of this study was to investigate the relationship between individual target tracking and force matching ability in a VR environment and hand smoothness in a simulated LVH repair. The results of this study show significant relationships between an individual's VR motor skills and their hand smoothness in an LVH repair. "Translational Smoothness of Dominant Hand" was found to be the most informative metric in this research as it has significant negative correlations to all of the target tracking and force matching abilities within the VR environment. The findings have great potential to reveal underlying shared psychomotor skill between the LVH procedure and force matching and tracking abilities in the defined VR environment. By investigating and validating the use of VR as a reliable source of motion metric evaluation, we hope to contribute to the development of objective and unbiased assessment techniques in the evaluation of the LVH repair procedure to improve LVH surgical education.

APPENDIX

Table III represents the details of the force matching module. There were 20 trials in this module for each participant that are selected from 16 Trial Conditions (TC). For instance, if "F1" is assigned to a participant, that participant faced one trial from Trial Condition 1, two trials from Trial Condition 7, and three trials from Trial Condition 9 (and so on). "F1", "F2", and "F3" were assigned to the participants randomly to 13, 12, and 9 individuals respectively. Table III represents the time limit (TL) (s), force magnitude (FM) (N), and force slope (FS) (N/s), which was defined as the speed at which the applied force increased from zero to the corresponding FM. The force direction is shown with the unit vector values of the axes X, Y, and Z, Fig. 4.

TABLE III. SUMMARY OF TRIAL CONDITION PARAMETERS IN RUNNING FORCE MATCHING MODULE IN VR

TC	TL (s)	FM (N)	FS (N/s)	Force Direction X, Y, Z	F1	F2	F3
1	10	3	3	0.58, 0.58, 0.58	1	1	1
2	7	3	10	-0.27, 0.80, -0.53	0	1	0
3	7	2	15	-0.77, -0.15, 0.62	1	0	1
4	7	3	5	0.37, -0.18, -0.91	0	1	2
5	7	3	12	-0.89, 0.15, -0.44	3	2	1
6	7	2	10	-0.24, -0.24, -0.94	1	0	1
7	7	4	10	-0.33, 0.67, -0.67	2	2	2
8	7	5	15	-0.87, 0.44, 0.22	2	2	2
9	7	4.5	5	0.80, 0.53, 0.27	3	3	3
10	7	4	7	-0.89, 0.15, 0.44	1	2	1
11	7	4.5	10	-0.30, -0.30, 0.91	2	1	1
12	7	5	10	-0.23, -0.69, 0.69	1	2	2
13	7	5.5	3	-0.33, -0.67, -0.67	1	1	1
14	7	6.5	2	-0.87, -0.44, -0.22	0	1	1
15	7	5.5	3	-0.89, -0.15, -0.44	1	0	0
16	7	6.5	3	-0.30, -0.30, -0.91	1	1	1

Likewise, the setting of the target tracking module is tabulated in Table IV. There were 15 trials in this module selected from 12 TC. Table IV also demonstrates angular speed (AS) (rad/s) and rotational direction (RD). The axes X, Y, and Z, of the force direction are shown in Fig. 5. In a random selection, "T1", "T2", and "T3" were assigned to 13, 12, and 9 participants respectively.

TABLE IV. SUMMARY OF TRIAL CONDITION PARAMETERS IN RUNNING TARGET TRACKING MODULE IN VR

TC	TL (s)	FM (N)	AS-RD (rad/s)	Force Direction X, Y, Z	T1	T2	T3
1	10	2	0.314-CCW	0.58, 0.58, 0.58	1	1	1
2	13	4	0.628-CW	0.80, -0.27, -0.53	1	1	1
3	8	3	0.628-CCW	0.51, 0.69, -0.51	1	1	1
4	5	4	0.314-CW	0.18, -0.91, -0.37	1	1	1
5	13	5	0.628-CCW	0.80. -0.27, 0.53	2	2	2
6	9	5.5	0.628-CW	0.64, 0.43, -0.64	0	1	1
7	8	4	0.5024-CCW	0.58, 0.58, 0.58	3	1	2
8	9	5.5	3.768 - CW	0.80, -0.27, 0.53	1	1	1
9	5	6	0.7536-CCW	0.51, 0.69, -0.51	1	1	1
10	10	3	0.4396-CW	0.18, -0.91, 0.37	2	1	2
11	13	5	6.28-CCW	0.49, 0.81, 0.32	1	1	1
12	10	5.5	0.942-CW	0.58, -0.58, 0.58	1	2	1

ACKNOWLEDGMENT

Funding for this study came from the Department of Defense grant #W81XWH-13-1-0080 entitled "Psycho-Motor and Error Enabled Simulations Modeling Vulnerable Skills in the Pre-Mastery Phase – Medical Practice Initiative Procedural Skill Decay and Maintenance (MPI-PSD)". Also, the authors thank Drew Rutherford with the Department of Health Professions, College of Science and Health, University of Wisconsin-La Crosse, WI 54601, USA.

REFERENCES

[1] E. P. Misiakos, P. Patapis, N. Zavras, P. Tzanetis, and A. Machairas, "Current Trends in Laparoscopic Ventral Hernia Repair," Jsls, 3, Miami, FL, 2015.

[2] M. van't Riet, E. W. Steyerberg, J. Nellensteyn, H. J. Bonjer, and J. Jeekel, "Meta-analysis of techniques for closure of midline abdominal incisions," British Journal of Surgery, vol. 89, no. 11, pp. 1365-2168, 2002.

[3] R. W. Luijendijk, W. C. J. Hop, M. P. van den Tol, D. C. D. de Lange, M. M. J. Braaksma, J. N. M. Ijzermans, R. U. Boelhouwer, B. C. de Vries, M. K. M. Salu, J. C. J. Wereldsma, C. M. A. Bruijninckx, and J. Jeekel, “A Comparison of Suture Repair with Mesh Repair for Incisional Hernia,” New England Journal of Medicine, vol. 343, no. 6, pp. 392-398, 2000.

[4] D. H. Shell Iv, J. de la Torre, P. Andrades, and L. O. Vasconez, “Open Repair of Ventral Incisional Hernias,” Surgical Clinics of North America, vol. 88, no. 1, pp. 61-83, 2008.

[5] R. Reznick, G. Regehr, H. MacRae, J. Martin, and W. McCulloch, “Testing technical skill via an innovative “bench station” examination,” The American Journal of Surgery, vol. 173, no. 3, pp. 226-230, 1997.

[6] J. F. Bodle, S. J. Kaufmann, D. Bisson, B. Nathanson, and D. M. Binney, “Value and face validity of objective structured assessment of technical skills (OSATS) for work based assessment of surgical skills in obstetrics and gynaecology,” Medical Teacher, vol. 30, no. 2, pp. 212-216, 2008.

[7] E. Hiemstra, W. Kolkman, R. Wolterbeek, B. Trimbos, and F. W. Jansen, “Value of an objective assessment tool in the operating room,” Canadian Journal of Surgery, vol. 54, no. 2, pp. 116-122, 2011.

[8] A.-L. D. D'Angelo, E. R. Cohen, C. Kwan, S. Laufer, C. Greenberg, J. Greenberg, D. Wiegmann, and C. M. Pugh, “Use of decision-based simulations to assess resident readiness for operative independence,” The American Journal of Surgery, vol. 209, no. 1, pp. 132-139, 2015.

[9] I. Oropesa, P. Sánchez-González, P. Lamata, M. K. Chmarra, J. B. Pagador, J. A. Sánchez-Margallo, F. M. Sánchez-Margallo, and E. J. Gómez, “Methods and Tools for Objective Assessment of Psychomotor Skills in Laparoscopic Surgery,” Journal of Surgical Research, vol. 171, no. 1, pp. e81-e95, 2011.

[10] A.-L. D. D'Angelo, D. N. Rutherford, R. D. Ray, S. Laufer, A. Mason, and C. M. Pugh, “Working volume: validity evidence for a motion-based metric of surgical efficiency,” The American Journal of Surgery, vol. 211, no. 2, pp. 445-450, 2016.

[11] A.-L. D. D'Angelo, D. N. Rutherford, R. D. Ray, S. Laufer, C. Kwan, E. R. Cohen, A. Mason, and C. M. Pugh, “Idle time: an underdeveloped performance metric for assessing surgical skill,” The American Journal of Surgery, vol. 209, no. 4, pp. 645-651, 2015.

[12] I. Oropesa, P. Sánchez-González, M. K. Chmarra, P. Lamata, A. Fernández, J. A. Sánchez-Margallo, F. W. Jansen, J. Dankelman, F. M. Sánchez-Margallo, and E. J. Gómez, “EVA: Laparoscopic Instrument Tracking Based on Endoscopic Video Analysis for Psychomotor Skills Assessment,” Surgical Endoscopy, vol. 27, no. 3, pp. 1029-1039, 2012.

[13] I. Oropesa, P. Sánchez-González, M. K. Chmarra, P. Lamata, R. Pérez-Rodríguez, F. W. Jansen, J. Dankelman, and E. J. Gómez, “Supervised classification of psychomotor competence in minimally invasive surgery based on instruments motion analysis,” Surgical Endoscopy, vol. 28, no. 2, pp. 657-670, 2013.

[14] E. F. Hofstad, C. Våpenstad, M. K. Chmarra, T. Langø, E. Kuhry, and R. Mårvik, “A study of psychomotor skills in minimally invasive surgery: what differentiates expert and nonexpert performance,” Surgical Endoscopy, vol. 27, no. 3, pp. 854-863, 2012.

[15] K. R. Sickle, D. A. McClusky, A. G. Gallagher, and C. D. Smith, “Construct validation of the ProMIS simulator using a novel laparoscopic suturing task,” Surgical Endoscopy And Other Interventional Techniques, vol. 19, no. 9, pp. 1227-1231, 2005.

[16] M. G. C. Pellen, L. F. Horgan, J. R. Barton, and S. E. Attwood, “Construct validity of the ProMIS laparoscopic simulator,” Surgical Endoscopy, vol. 23, no. 1, pp. 130-139, 2008.

[17] H. Mohamadipanah, C. Parthiban, L. Maulson, S. Laufer, D. Rutherford, K. Law, S. DiMarco, and C. Pugh, “Is Indecisiveness Linked to Hand Movements in Performing Laparoscopic Ventral Hernia Repair?,” in 1th International Workshop on Surgical Data Science, Heidelberg, Germany, 2016.

[18] C. Pugh, S. Plachta, E. Auyang, A. Pryor, and E. Hungness, “Outcome measures for surgical simulators: Is the focus on technical skills the best approach?,” Surgery, vol. 147, no. 5, pp. 646-654, 2010.

[19] A.-L. D. D'Angelo, K. E. Law, E. R. Cohen, J. A. Greenberg, C. Kwan, C. Greenberg, D. A. Wiegmann, and C. M. Pugh, “The use of error analysis to assess resident performance,” Surgery, vol. 158, no. 5, pp. 1408-1414, 2015.

[20] H. Mohamadipanah, C. T. Millar, D. N. Rutherford, K. E. Law, R. D. Ray, and C. M. Pugh. "How Motion Tracking Relates to End Product Quality in Laparoscopic Procedures," 2/18/2016; http://www.asc-abstracts.org/abs2016/48-05-how-motion-tracking-relates-to-end-product-quality-in-laparoscopic-procedures/.

[21] V. Falk, “Manual control and tracking—a human factor analysis relevant for beating heart surgery,” The Annals of Thoracic Surgery, vol. 74, no. 2, pp. 624-628, 2002.

[22] N. Hogan, and D. Sternad, “Sensitivity of Smoothness Measures to Movement Duration, Amplitude and Arrests,” Journal of motor behavior, vol. 41, no. 6, pp. 529-534, 2009.

Is Indecisiveness Linked to Hand Movements in Performing Laparoscopic Ventral Hernia Repair?

Conference Paper · June 2016

8 authors, including:


Hossein Mohamadipanah
University of Wisconsin–Madison


Shlomi Laufer
University of Wisconsin–Madison


Drew N Rutherford
University of Wisconsin - La Crosse


Katherine Law Forsyth
Mayo Clinic - Rochester

Is Indecisiveness Linked to Hand Movements in Performing Laparoscopic Ventral Hernia Repair?

Hossein Mohamadipanah[1], Chembian Parthiban[2], Lakita Maulson[1], Shlomi Laufer[1], Drew Rutherford[3], Katherine Law[4], Shannon DiMarco[1], Carla Pugh[1,4]

[1]Department of Surgery, [2]Department of Mechanical Engineering, [3]Department of Kinesiology, and [4]Department of Industrial and Systems Engineering of the University of Wisconsin-Madison, WI, USA

Abstract. This study aims to ascertain whether indecisiveness is linked to fine motor skills during a LVH repair procedure. General surgery residents (N=34) from seven different programs throughout the Midwest were asked to participate in a Virtual Reality (VR) force discrimination task to make a decision within the defined time frame. Additionally, they were asked to perform a simulated Laparoscopic Ventral Hernia (LVH) repair in which the participants' hand movements were tracked using electromagnetic sensors. It was found that long periods of indecisiveness in discriminating forces within the VR environment showed a significant positive correlation ($p < 0.05$) with longer translational and angular path lengths of the participants' dominant hand while performing LVH repair. Long periods of indecisiveness were also positively correlated with greater working volumes when repairing a ventral hernia ($p < 0.05$). These findings indicate the potential for motion metrics to be used in objective assessment of laparoscopic skills. In addition, more work is needed to fully explore the relationship between indecisiveness and overall operative skill.

1 Introduction

Identifying objective means of assessment for surgical procedures has been a significant focus of research over the past decade [1-3]. With the emergence of new technologies, particularly virtual reality (VR) advancements, educating and assessing surgeon psychomotor skills outside of the operating room is actually plausible[4]. In fact, technology-based assessments have the potential to be more reliable and unbiased than human-based assessments, which are heavily observation dependent. Evaluation of psychomotor surgical skills via VR has been accomplished through hand motion metrics [5]. For instance, two motion metrics, smoothness and path length, have been used as objective assessment metrics in evaluating a set of pre-defined tasks on a laparoscopic VR simulator [6]. Similarly, a laparoscopic VR simulator was validated using an angular path length motion metric [7]. Although previous studies have evidence for the validity of certain motion metrics as objective measures of a surgeon's psychomotor skills, more explicit guidelines for training and remediation are needed in order to improve surgical practice. Better defining the relationship between the objective motion measures and specific medical procedures may help to bridge the gap. The purpose of this study was to identify hand motion metrics which

can play a role in the objective assessment of performance during a Laparoscopic Ventral Hernia (LVH) repair procedure. We aimed to investigate whether indecisiveness in VR was linked to the hand movements during a LVH repair procedure. Subsequently, the relationship between the motion metrics for the LVH repair procedure and the participant's ability to make a decision during a timed virtual task was the focus of this study. We hypothesized that there would be a significant, positive correlation between a participant's indecisiveness in discriminating stiffness in a VR setting and their psychomotor skills during a LVH repair procedure.

2 Method

2.1 Participants and Motion Tracking System

Thirty-four general surgery residents (PGY 1-4) were recruited on a voluntary-basis from seven different institutions across the Midwest to participate in this study. Each data collection was performed at the participant's designated institution either in a dedicated simulation or research space. When asked to identify hand dominance, 31 participants identified as right-handed, 2 participants identified as left-handed, and 1 participant identified as ambidextrous. For the purposes of this study, the ambidextrous participant was assumed to be right-handed. This study was approved by the University of Wisconsin – Madison Institutional Review Board.

Each participant was equipped with a head camera which recorded performance from their point of view. Additionally, motion tracking sensors (trackSTAR Model 180) from a MotionMonitor® System (Innovative Sports Training, Inc.) were affixed to the participants' index fingers, thumbs and wrists. These sensors provided six degrees of freedom, which include translational and angular motions. To remove noise, the motion data received by the sensors is filtered by a low-pass second-order Butterworth filter with a cutoff frequency of 1 Hz [8]. A total of four metrics of interest were utilized from the hand motion data while participants performed the LVH repair. The first metric, "Translational Path Length", is defined as the total traveled distance of each sensor throughout task execution. "Translational Working Volume", the second metric, is defined as each sensor's summation of distances from its average spatial location [9]. The third and fourth metrics, "Angular Path Length" and "Angular Working Volume", are equivalent to the first two, respectively but are defined using angular motions instead of translational motions.

2.2 Laparoscopic Ventral Hernia Simulator

The LVH simulator, shown in Fig. 1, was previously developed and validated [10, 11]. The simulator includes an abdominal box with a skin attachment that has a 10 cm diameter ventral hernia injury. For the purpose of this study, the laparoscopic ports were already placed for the participant prior to the start of the procedure. Additionally, two sets of transfascial sutures were pulled through the abdomen for the participants with the expectation for the participant to bring through, tie, and cut the final

two sets of transfascial suture as well as place the first five tacks in the allotted time (15 minutes). Except for cautery instruments, participants were provided with the same instruments necessary to complete the full LVH repair within 15 minutes. For this study, motion data analysis was performed only during the time it took for participants to secure and tie the first suture set.

Fig. 1. (Left) The outside view shows a participant performing the LVH procedure while the sensors are affixed to their hands, (Right) The inside view of the simulated patient's body seen through the endoscopic camera which participants visualize via a monitor.

2.3 Decision Making in Virtual Reality

Participants were also asked to perform a stiffness discrimination task via the use of a haptic device (Force Dimension Omega ®), shown in Fig. 2 (Left) which provides force feedback. The virtual environment includes a virtual stylus which moves within the 3D environment and provides force reactions when two predefined blocks, A and B, are tapped, shown in Fig. 2 (Right). Participants had to determine which stiffness was greater: block A and B. Within each condition, the stiffness of A or B were set between 10% and 90% of the maximum closed loop force handled by the device (14.5 N/mm), or between 1.45 and 13.05 N/mm. The effective stiffness difference between A and B was set between 0 and 8.7 N/mm, with 0 signifying equal magnitudes between A and B. Each participant faced 19 trials with each trial lasting 15, 20 or 30 seconds. The initial two trials were set for 30 seconds, so that each participant had a chance to understand the task at hand. Procedural-wise, the participants were asked to first differentiate the stiffness associated with blocks A and B. They were then expected to indicate which block had the higher stiffness by placing the virtual stylus over the respective block, then clicking the button on the haptic device stylus to finalize their decision. If participants believed the forces to be equal, they were instructed to click on a designated area between the two blocks. A timer was displayed on the screen letting participants know how many seconds were left in the trial. For all trials, total number of correct force discrimination is defined as “Correct Discrimination”, and similarly total number of incorrect force discrimination is defined as “Wrong Discrimination”, and finally “Time Out” describes the total number of times that a participant ran out of time prior to finalizing their decision.

Fig. 2. Stiffness differentiation; (Left) A participant operating the haptic device, (Right) Screen shot of the stiffness discrimination task in the virtual environment, where blocks A and B are different stiffness being rendered.

3 Results

The motion metrics were calculated for both the dominant and non-dominant hand separately. Each calculation included the average of the motion metrics provided by the three sensors placed on the dominant or non-dominant hands during the LVH repair. Correlation coefficients (Spearman's rho) were then calculated to explain the relationships between the translational and angular Path Lengths and Working Volumes metrics from the LVH repair, and the VR force discrimination metrics “Correct Discrimination”, “Wrong Discrimination”, and “Time Out”. Table 1 illustrates the correlation coefficients between the translational metrics of the LVH repair and the VR environment metrics. Higher “Time Out” was shown to have a significant positive correlation ($p < 0.05$) with two of the LVH repair metrics, “Translational Path Length - Dominant Hand” and “Translational Working Volume - Dominant Hand”. “Correct Discrimination” and “Wrong Discrimination” did not show a significant correlation ($p > 0.05$) to any translational metrics of the LVH repair procedure.

Table 1. The correlations between the translational metrics in LVH repair and VR force discrimination metrics.

Translational Metric of LVH Repair	***Metric of Stiffness Discrimination***	***R***	***P value***
Translational Path Length Non-Dominant Hand	Correct Discrimination	-0.1511	0.3936
	Wrong Discrimination	+0.0995	0.5756
	Time Out	+0.2277	0.1953
Translational Path Length Dominant Hand	Correct Discrimination	-0.3335	0.0539
	Wrong Discrimination	+0.2382	0.1749
	Time Out	**+0.3708**	**0.0309**
Translational Working Volume Non-Dominant Hand	Correct Discrimination	-0.1454	0.4118
	Wrong Discrimination	+0.1086	0.5411
	Time Out	+0.1399	0.4302
Translational Working Volume Dominant Hand	Correct Discrimination	-0.2437	0.1649
	Wrong Discrimination	+0.1196	0.5003
	Time Out	**+0.4640**	**0.0057**

In the same way, Table 2 shows the correlation coefficients between the VR environment metrics and the angular metrics of the LVH repair. Higher “Time Out” was shown to have a significant positive correlation ($p < 0.05$) with the LVH repair metric, “Angular Path Length – Dominant Hand”. Meanwhile, “Correct Discrimination” and “Wrong Discrimination” did not significantly correlate ($p > 0.05$) to any angular metrics from the LVH repair simulation.

Table 2. The correlations between the angular metrics in LVH repair and VR force discrimination metrics.

Angular Metric of LVH Repair	***Metric of Stiffness Discrimination***	***R***	***P value***
Angular Path Length Non-Dominant Hand	Correct Discrimination	-0.0761	0.6688
	Wrong Discrimination	+0.0181	0.9189
	Time Out	+0.2027	0.2502
Angular Path Length Dominant Hand	Correct Discrimination	-0.3273	0.0588
	Wrong Discrimination	+0.2484	0.1567
	Time Out	**+0.3740**	**0.0293**
Angular Working Volume Non-Dominant Hand	Correct Discrimination	+0.1442	0.4158
	Wrong Discrimination	-0.2105	0.2321
	Time Out	+0.1561	0.3780
Angular Working Volume Dominant Hand	Correct Discrimination	-0.1162	0.5127
	Wrong Discrimination	+0.0880	0.6208
	Time Out	+0.2222	0.2065

4 Discussion

In this study, we investigated Path Length and Working Volume as two hand motion metrics that can potentially play a role in the objective assessment of the psychomotor skills necessary to successfully complete a LVH repair procedure. Our results, shown in Table 1 and Table 2, indicate three motion metrics that have shown significant correlation to stiffness differentiation metrics in VR: 1. “Translational Path Length - Dominant Hand”; 2. “Translational Working Volume - Dominant Hand”, and; 3. “Angular Path Length - Dominant Hand”. It is also worth mentioning that non-dominant hand metrics did not correlate to any metrics in stiffness differentiation. More importantly the findings show that higher values of these three metrics have significant positive correlations to the VR environment metric “Time Out”. Therefore, we were able to show that participants who had difficulty discriminating between reaction forces within a given time limit, also had greater dominant hand translational and angular Path Lengths and had greater dominant hand angular Working Volumes.
Further work need to be done as the generalizability of this study is limited by the small sample size and participant exclusion due to volunteer-dependent recruitment. Future work will include a greater number of participants, both experts and surgical residents. In general, the findings are notable because the current literature shows that

greater Path Length and Working Volume usually tend to correlate with lower procedure performance [6, 9]. Therefore, we aim to further our findings in the future and investigate the relationship between indecisiveness in VR and global LVH performance ratings obtained by experts. This work will facilitate the development of automated, objective performance metrics that will supplement faculty feedback and potentially decrease the time needed to conduct valid and reliable performance assessments.

References

1. A.-L. D. D'Angelo, D. N. Rutherford, R. D. Ray, S. Laufer, C. Kwan, E. R. Cohen, A. Mason, and C. M. Pugh, "Idle time: an underdeveloped performance metric for assessing surgical skill," *The American Journal of Surgery,* vol. 209, no. 4, pp. 645-651, 2015.
2. A. Dosis, R. Aggarwal, F. Bello, K. Moorthy, Y. Munz, D. Gillies, and A. Darzi, "Synchronized video and motion analysis for the assessment of procedures in the operating theater," *Archives of Surgery,* vol. 140, no. 3, pp. 293-299, 2005.
3. V. Datta, M. Mandalia, S. Mackay, A. Chang, N. Cheshire, and A. Darzi, "Relationship between skill and outcome in the laboratory-based model," *Surgery,* vol. 131, no. 3, pp. 318-323, 2002.
4. R. Aggarwal, K. Moorthy, and A. Darzi, "Laparoscopic skills training and assessment," *British Journal of Surgery,* vol. 91, no. 12, pp. 1549-1558, 2004.
5. I. Oropesa, P. Sánchez-González, P. Lamata, M. K. Chmarra, J. B. Pagador, J. A. Sánchez-Margallo, F. M. Sánchez-Margallo, and E. J. Gómez, "Methods and Tools for Objective Assessment of Psychomotor Skills in Laparoscopic Surgery," *Journal of Surgical Research,* vol. 171, no. 1, pp. e81-e95, 2011.
6. M. G. C. Pellen, L. F. Horgan, J. R. Barton, and S. E. Attwood, "Construct validity of the ProMIS laparoscopic simulator," *Surgical Endoscopy,* vol. 23, no. 1, pp. 130-139, 2008.
7. K. W. Dongen, E. Tournoij, D. C. Zee, M. P. Schijven, and I. A. M. J. Broeders, "Construct validity of the LapSim: Can the LapSim virtual reality simulator distinguish between novices and experts?," *Surgical Endoscopy,* vol. 21, no. 8, pp. 1413-1417, 2007.
8. V. Falk, "Manual control and tracking—a human factor analysis relevant for beating heart surgery," *The Annals of Thoracic Surgery,* vol. 74, no. 2, pp. 624-628, 2002.
9. A.-L. D. D'Angelo, D. N. Rutherford, R. D. Ray, S. Laufer, A. Mason, and C. M. Pugh, "Working volume: validity evidence for a motion-based metric of surgical efficiency," *The American Journal of Surgery,* vol. 211, no. 2, pp. 445-450, 2016.
10. A.-L. D. D'Angelo, K. E. Law, E. R. Cohen, J. A. Greenberg, C. Kwan, C. Greenberg, D. A. Wiegmann, and C. M. Pugh, "The use of error analysis to assess resident performance," *Surgery,* vol. 158, no. 5, pp. 1408-1414, 2015.
11. C. Pugh, S. Plachta, E. Auyang, A. Pryor, and E. Hungness, "Outcome measures for surgical simulators: Is the focus on technical skills the best approach?," *Surgery,* vol. 147, no. 5, pp. 646-654, 2010.

JOURNAL OF SURGICAL RESEARCH • SEPTEMBER 2016 (205) 192–197

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.JournalofSurgicalResearch.com

Association for Academic Surgery

Exploring hand coordination as a measure of surgical skill

Katherine E. Law, MS,[a] *Caitlin G. Jenewein, BS,*[b] *Samantha J. Gannon, BS,*[b] *Shannon M. DiMarco, BA,*[b] *Lakita J. Maulson, BS,*[b] *Shlomi Laufer, PhD,*[b] *and Carla M. Pugh, MD, PhD*[a,b,*]

[a] *Department of Industrial and Systems Engineering, University of Wisconsin–Madison, Madison, Wisconsin*

[b] *Department of Surgery, School of Medicine and Public Health, University of Wisconsin–Madison, Madison, Wisconsin*

ARTICLE INFO

Article history:
Received 5 February 2016
Received in revised form
3 May 2016
Accepted 10 June 2016
Available online 17 June 2016

Keywords:
Coordination
Motor movement
Suture grasping
Laparoscopic skills

ABSTRACT

Background: The study aim was to identify residents' coordination between dominant and nondominant hands while grasping for sutures in a laparoscopic ventral hernia repair procedure simulation. We hypothesize residents will rely on their dominant and nondominant hands unequally while grasping for suture.

Methods: Surgical residents had 15 min to complete the mesh securing and mesh tacking steps of a laparoscopic ventral hernia repair procedure. Procedure videos were coded for manual coordination events during the active suture grasping phase. Manual coordination events were defined as: active motion of dominant, nondominant, or both hands; and bimanual or unimanual manipulation of hands. A chi-square test was used to discriminate between coordination choices.

Results: Thirty-six residents (postgraduate year, 1-5) participated in the study. Residents changed manual coordination types during active suture grasping 500 times, ranging between 5 and 24 events ($M = 13.9$ events, standard deviation [SD] $= 4.4$). Bimanual coordination was used most (40%) and required the most time on average ($M = 20.6$ s, $SD = 27.2$), while unimanual nondominant coordination was used least (2.2%; $M = 7.9$ s, $SD = 6.9$). Residents relied on their dominant and nondominant hands unequally ($P < 0.001$). During 24% of events, residents depended on their nondominant hand ($n = 120$), which was predominantly used to operate the suture passer device.

Conclusions: Residents appeared to actively coordinate both dominant and nondominant hands almost half of the time to complete suture grasping. Bimanual task durations took longer than other tasks on average suggesting these tasks were characteristically longer or switching hands required a greater degree of coordination. Future work is necessary to understand how task completion time and overall performance are affected by residents' hand utilization and switching between dominant and nondominant hands in surgical tasks.

* *Corresponding author*. Department of Surgery, University of Wisconsin Hospital and Clinics, 600 Highland Avenue, K6/120 CSC, Madison, WI 53792. Tel.: +1 (608) 263 5218; fax: (608) 263 2354.

E-mail address: pugh@surgery.wisc.edu (C.M. Pugh).

0022-4804/$ – see front matter
http://dx.doi.org/10.1016/j.jss.2016.06.038

Introduction

Advanced manual coordination for manipulating objects is critical when performing tasks that require fine motor movement. Tasks become more challenging when performed during open surgery or laparoscopically, when instruments act as extensions of the hands and tactile feedback is significantly reduced.[1] Motor movement can be separated into unimanual or bimanual coordination tasks, depending on whether the movement is made using one or both hands, respectively. In surgical procedures, movements that use bimanual coordination of both of hands are common to complete steps such as dissection or suturing.[2] Although early movement research identified the implications of various performance measures of unimanual movement, focus has shifted toward the implications of bimanual coordination of multiple movements.[3]

Although most coordination movements generally involve the use of two hands, we typically develop a greater reliance on one hand over the other. This dominant hand generally carries out tasks that require a more advanced level of accuracy. As a result, the dominant hand can develop enhanced skill and accuracy in performing various motor behaviors when analyzing metrics such as amplitude, force, and time.[4,5]

Similar findings were shown in medical students whose working volumes were smaller for their dominant hand while performing a suturing task.[6] These findings support the role hand dominance can play in the overall function of coordination and the control of bimanual movements.

With technological innovation, analysis of surgical hand movement is growing. Those with advanced dexterity are considered to have rehearsed and mastered such surgical skill.[7] Furthermore, motion metrics are used to understand and compare novice surgeons to experts.[6,8] With various approaches available to measure surgical and laparoscopic skill, some suggest a more comprehensive assessment should be used.[9] Combining clinical, technological, and analytical aspects of surgical ability together into one larger assessment could provide a more accurate picture of surgical ability. As a first step toward reaching comprehensive assessment, the aim of our study was to understand coordination of residents' dominant and nondominant hands during a laparoscopic surgical simulation. Specifically, we will look at residents' unimanual and bimanual motor coordination while grasping for suture to complete a simulated laparoscopic ventral hernia (LVH) repair. We hypothesize residents will rely on their dominant and nondominant hands unequally.

Methods

Setting and participants

Residents (postgraduate year [PGY], 1-5) from seven general surgery training programs located in the Midwest participated in this study. Over a period of 3 mo, a data collection for a larger study was conducted at each participating institution to maximize opportunity for resident participation. Although the data collections for the larger study specifically sought out residents in research years, both research and clinical residents participated. The University of Wisconsin Hospitals and Clinics Institutional Review Board approved this study.

Research protocol

Before the simulated procedure, residents completed a demographic survey where hand dominance was self-reported.[10] Only residents that reported as right-hand or left-hand dominant were included in this analysis, with those identifying as ambidextrous or multihanded excluded.

Residents were expected to complete the mesh securing and mesh tacking steps of a simulated LVH repair procedure within 15 min. The validated LVH simulator[11] contained a midline, 10 × 10-cm ventral hernia located 5 cm above the umbilicus with mesh partially attached to the abdomen with transfascial sutures. Due to the advanced skill required to complete an LVH repair procedure, residents were presented with a scenario where laparoscopic ports were placed, the hernia and mesh were measured, the mesh was appropriately sized, and two of four transfascial suture sets were already secured to the abdomen, as shown in Figure.

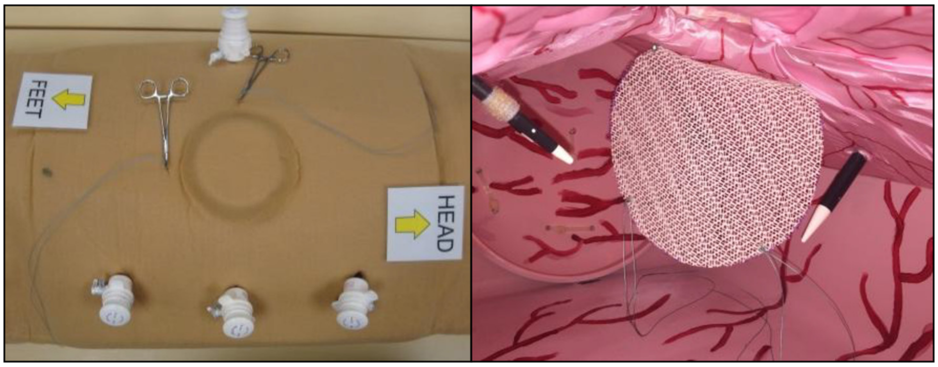


Figure – During a simulated LVH repair procedure, residents were presented with the simulated abdomen of a patient (left) that had two sets of transfascial suture already brought up (right). Residents were expected to secure two sets of transfascial sutures to the abdominal wall before tacking. (Color version of figure is available online).

One researcher acted as the resident's surgical assistant. The surgical assistant was considered at the skill level of a medical student, and their duties included holding and manipulating the endoscopic camera, and limited tasks such as holding the suture passer. Audio and video of the surgical area and internal endoscopic video were recorded during each resident's procedure. Residents were expected to bring up and secure the last two sets of transfascial sutures (i.e., mesh securing) and place five tacks to secure the mesh to the abdominal wall (i.e., mesh tacking). Residents were provided with all laparoscopic instruments required to perform a full LVH repair, except for cautery instruments.

After data collection, audio and video recordings of each repair were reviewed and coded using Transana transcription software (University of Wisconsin–Madison). Three phases of the procedure were identified: hernia visualization, mesh securing, and mesh tacking. This study only analyzed the manual coordination used for grasping suture during the mesh securing step. Manual coordination frequency and duration were identified, and the grasp attempt frequency required to complete the mesh securing step of the repair was tracked for each resident.

Manual coordination events and analysis

Manual coordination events were categorized based on two characteristics: (1) dominant or nondominant hand use and (2) manual movement type. Manual coordination events were tagged for active motion of the resident's dominant, nondominant, or both hands, as defined in Table 1. In addition, coordination events were coded for bimanual or unimanual manipulation of instruments during suture grasping attempts, as defined in Table 1. Bimanual movement involved using both hands actively or one hand actively with another resting on equipment or instruments for the repair (e.g., statically holding laparoscopic grasper). In contrast, unimanual movement involved using only one hand actively with the other hand stationary and outside the surgical area (e.g., at the resident's side).

Instruments and equipment used by residents were also identified during each event in order quantify active time spent using each instrument. All tagged codes along with their respective time segments were exported to IBM SPSS Statistics version 23 (IBM Corp, Armonk, NY) for descriptive and statistical analyses. Chi-square analyses were performed to compare resident reliance on their dominant and nondominant hands. A one-way analysis of variance was performed to compare coordination event durations across the event types.

Suture grasping attempts

While completing video tagging for coordination events, some residents were observed as having difficulty successfully grasping suture or transferring suture between a laparoscopic grasper and a suture passer. As a result, each active attempt to grasp suture was tallied as residents completed the procedure.

Results

Thirty-six residents (52.8% female and 91.7% right-handed) participated in the study and met inclusion criteria for this analysis. One resident opted-out of the study and nine residents identified as ambidextrous or did not specify hand dominance and were therefore excluded from this analysis. The final group of residents consisted of various experience levels ($n = 36$; PGY1 = 3, PGY2 = 12, PGY3 = 13, PGY4 = 6, and PGY5 = 2).

Residents transitioned between various types of manual coordination during active suture grasping 500 times in total, with coordination movements ranging between 5 and 24 events per resident (M = 13.9 events and standard deviation [SD] = 4.4; see Table 2).

The most common type of coordination during suture grasping was bimanual events with both hands ($n = 200$; M = 6.1 events, SD = 3.4). These events required significantly more time (M = 20.6 s, SD = 27.2 s) in comparison to other event types ($F(3,196) = 4.366$, $P = 0.005$). Bimanual events with both hands commonly involved coordinating a suture passer and laparoscopic grasper (87.0%) to pass the suture from the grasper to the suture passer to pull it through the abdominal wall, as summarized in Table 3. The second most frequently used equipment involved laparoscopic graspers, typically

Table 1 – Definitions and codes of manual coordination events identified during active attempts to grasp suture during the mesh-securing phase of a simulated laparoscopic ventral hernia repair procedure.

Keyword	Code	Definition
Bimanual movement		
Bimanual: both moving	BA	Both hands on instruments or simulated patient, actively moving
Bimanual: dominant hand active, nondominant hand static	BD	Dominant hand actively moving, with nondominant hand stationary on an instrument or simulated patient
Bimanual: dominant hand static, nondominant hand active	BN	Nondominant hand actively moving, with dominant hand stationary on an instrument or simulated patient
Unimanual movement		
Unimanual: dominant hand active	UD	Dominant hand active, while nondominant hand does not touch instruments or simulated patient
Unimanual: nondominant hand active	UN	Nondominant hand active, while dominant hand does not touch instruments or simulated patient

Table 2 – Residents' manual coordination events types, frequency, mean duration and range during suture grasping, and mean grasp attempts to complete repair.

Coordination events	Frequency (total = 500)	Mean per resident (SD)	Mean duration, s (SD)	Duration range, s	Mean grasp attempts (SD)
Bimanual: both active	200	6.1 (3.4)	20.6 (27.2)	1.5-65.1	4.3 (0.7)
Bimanual: dominant active, nondominant static	159	4.4 (3.1)	12.1 (9.1)	0.5-58.5	4.4 (1.0)
Bimanual: dominant static, nondominant active	109	3.0 (2.6)	10.1 (6.8)	0.8-30.5	4.3 (0.6)
Unimanual: dominant active	21	0.6 (1.5)	13.3 (12.8)	3.4-57.3	4.5 (0.5)
Unimanual: nondominant active	11	0.3 (1.2)	8.2 (6.6)	2.1-19.7	4.1 (0.3)

when residents moved suture to the suture passer with both graspers.

Unimanual events occurred 32 times in total, as these events occurred on average 0.9 times (SD = 1.4) per resident. While infrequent, residents typically used unimanual coordination when they successfully grasped the suture with the suture passer and pulled the suture through. In some circumstances, residents directed the surgical assistant to hold the camera and one laparoscopic grasper as they manipulated the suture passer in an attempt to grasp suture.

Residents did not rely on their dominant and nondominant hands equally (χ^2 [2, $n = 36$] = 33.27; $P < 0.001$).

Residents used their dominant hand a majority of the time, either primarily in bimanual events ($n = 159$ events, M = 4.4, SD = 3.1) or solely in unimanual events ($n = 21$ events, M = 0.6, SD = 1.5).

Only during 24% of all manual coordination events did residents depend on their nondominant hand ($n = 120$ events, M = 3.6, SD = 3.8). Furthermore, only four residents used unimanual coordination actively with their nondominant hand. In these cases, residents directed the surgical assistant to hold the suture passer stationary while the resident used their nondominant hand to transfer suture from a laparoscopic grasper to the suture passer.

Although actively grasping the transfascial sutures, residents took on average 4.25 attempts (SD = 0.87) to successfully secure the remainder of the mesh to the abdominal wall. While four residents did this step incorrectly and pulled both suture sets through the abdominal wall together, 13 residents (36%) required multiple attempts to grasp the suture with the suture passer.

Table 3 – Frequency and mean frequency per participant of instrument and equipment coordination during bimanual events using both hands actively.

Instrument/equipment combinations	Frequency (%), (total = 200)
Suture passer, laparoscopic grasper	174 (87.0)
Laparoscopic graspers	21 (10.5)
Suture passer*	4 (2.0)
Suture passer, endoscope	1 (0.5)
Laparoscopic grasper, suture passer	1 (0.0)

*In these instances, both hands used the suture passer.

Discussion

This study analyzed residents' motor performance during a simulated LVH repair procedure. Residents used various types of manual coordination methods during active suture grasping and required additional attempts to grasp transfascial suture to successfully secure the mesh over the hernia.

Residents appeared to actively coordinate both dominant and nondominant hands almost half of the time to complete suture grasping. Bimanual tasks that required active use of both hands frequently involved transferring suture between laparoscopic graspers and suture passers. The durations of bimanual tasks were statistically longer on average, suggesting it was characteristic of the task, or switching hands required a greater degree of coordination or cognitive load. We suspect the high occurrence of bimanual coordination with both hands in our present study may be reflective of resident inexperience. One study demonstrated efficient movement during suturing was reflective of experience and proficiency.[2] As a result, we anticipate this ability and experience would translate over to laparoscopy where laparoscopic attending surgeons would use more efficient movements.

Few residents relied on their nondominant hand primarily, which supports evidence that most individuals rely on one hand over the other.[12] With only four residents using their dominant hand as the primary hand in unimanual tasks, we speculate that those residents felt less confident in their surgical abilities when relying primarily on their nondominant hand. Because residents understood their surgical performance would be analyzed, we can speculate that they wanted to perform the procedure as proficiently as possible and thus most likely focused use on their dominant hand to diminish the risk of demonstrating poor surgical ability. Future surgical task practice in simulation should focus on the improvement of coordination for bimanual task performance and further develop nondominant hand use to improve dexterity.[7]

More than one-third of the residents required additional efforts to secure the transfascial suture. This may suggest laparoscopic grasping is a skill perfected with time. Through identifying manual coordination techniques and active attempts at grasping suture, we were also able to indirectly identify residents that had performed one step of a LVH repair incorrectly. By requiring residents to bring up and secure the final two suture sets from the mesh, four grasps at

a minimum number were expected—two for each suture pair. Most residents required the minimum four or more attempts. Four residents performed the step in three attempts or less which indicates that they did not secure the suture into the abdomen with a fascial bridge or were unable to complete the procedure. However, this metric alone does not indicate whether residents required multiple grasp attempts to bring up one suture at a time. Identifying excessive grasp attempts may be an initial step in developing coordination metrics to assess surgical skill. Evaluating grasp attempts and coordination (i.e., hand dominance and reliance, bimanual and unimanual movements) in expert laparoscopic surgeons will be necessary to set standards of proficiency in surgical residents.

There were several limitations to this study. First, this study allowed residents to perform a simulated grasping task as they would intraoperatively. Discriminating between manual coordination types may have been better compared using experimentally-designed studies; however, our results demonstrate the variable ways residents coordinated suture grasping. Second, residents were only required to complete the latter part of the LVH repair procedure. Requiring residents to insert and orient the mesh and bring up the first sets of transfascial sutures may provide further understanding of how residents negotiate the laparoscopic instruments and their surgical ability. Third, participating residents were equipped with motion monitoring readers on their fingers and wrists; however, that data were not available at the time of analysis and so video tagging was used alternatively. While definitions of each classification were determined and clarified throughout video tagging, the possibility of human error is inherent in this style of data collection. At last, only three residents identified as left-handed in this study, and these findings may be biased toward right-handed residents. Testing an evenly distributed population of right- and left-handed participants may reveal more generalizable conclusions to the larger surgical community.

This study demonstrates that identification of motor movements and hand dominance during surgical tasks may be one method of assessing effective coordination. Our results provide evidence that residents use multiple motor coordination methods, and some favor their dominant hand when suture grasping. Improved use of residents' nondominant hand may have the potential to improve bimanual coordination. Furthermore, skill development to transition to one-handed bimanual movements in which one hand is largely active while the other, while used, is mostly stationary may also improve success rates of manual surgical tasks. Future work is necessary to understand how task completion time and overall performance are affected by residents' hand utilization and switching between dominant and nondominant hands in surgical tasks. As this study presented a midline hernia scenario for analysis, further work should explore whether laterality of the hernia impacts the facility of either handedness. Finally, of the institutions whose residents participated in this study, some incorporated simulation in their curriculum and some did not. Further work is needed to determine if resident coordination and performance differs between institutions and their levels of simulation-based training.

Acknowledgment

Funding for this study was provided by the US Army Medical Research Acquisition Activity grant entitled "Psycho-Motor and Error Enabled Simulations: Modeling Vulnerable Skills in the Pre-Mastery Phase" W81XWH-13-1-008.

Authors' contributions: K.E.L. contributed to study conception and design; data collection, analysis and interpretation; writing the manuscript and critical revision of the manuscript. S.J.G. contributed to study conception and design; data collection, analysis and interpretation; writing the manuscript and critical revision of the manuscript. C.G.J. contributed to study data collection, writing the manuscript and critical revision of the manuscript. L.J.M. contributed to study data collection, writing the manuscript and critical revision of the manuscript. S.M.D. contributed to study data collection, writing the manuscript and critical revision of the manuscript. S.L. contributed to study data collection, writing the manuscript, and critical revision of the manuscript. C.M.P. contributed to study conception and design; data collection, analysis and interpretation; writing the manuscript and critical revision of the manuscript.

Disclosure

The authors reported no proprietary or commercial interest in any product mentioned or concept discussed in this article.

REFERENCES

1. Bonrath EM, Dedy NJ, Zevin B, Grantcharov TP. Defining technical errors in laparoscopic surgery: a systematic review. *Surg Endosc*. 2013;27:2678–2691.
2. D'Angelo AL, Rutherford DN, Ray RD, et al. Working volume: validity evidence for a motion-based metric of surgical efficiency. *Am J Surg*. 2016;211:445–450.
3. Swinnen SP, Wenderoth N. Two hands, one brain: cognitive neuroscience of bimanual skill. *Trends Cog Sci*. 2004;8:18–25.
4. Serrien DJ, Wiesendanger M. A higher-order mechanism overrules the automatic grip-load force constraint during bimanual asymmetrical movements. *Behav Brain Res*. 2001;118:153–160.
5. Franz EA, Zelaznik HN, McCabe G. Spatial topological constraints in a bimanual task. *Acta Psychol*. 1991;77:137–151.
6. D'Angelo AL, Rutherford DN, Ray RD, et al. Idle time: an underdeveloped performance metric for assessing surgical skill. *Am J Surg*. 2015;209:645–651.
7. Bann S, Khan MS, Datta VK, Darzi AW. Technical performance: relation between surgical dexterity and technical knowledge. *World J Surg*. 2004;28:142–146.
8. Oropesa I, Chmarra MK, Sánchez-González P, et al. Relevance of motion-related assessment metrics in laparoscopic surgery. *Surg Innovation*. 2013;20:299–312.

9. Oropesa I, Sánchez-González P, Lamata P, et al. Methods and tools for objective assessment of psychomotor skills in laparoscopic surgery. *J Surg Res*. 2011;171:e81–e95.
10. D'Angelo AL, Ray RD, Jenewein CG, Jones GF, Pugh CM. Residents' perception of skill decay during dedicated research time. *J Surg Res*. 2015;199:23–31.
11. Pugh CM, Plachta S, Auyang E, Pryor A, Hungness E. Outcome measures for surgical simulators: is the focus on technical skills the best approach? *Surgery*. 2010;147:646–654.
12. Badalato GM, Shapiro E, Rothberg MB, et al. The da Vinci robot system eliminates multispecialty surgical trainees' hand dominance in open and robotic surgical settings. *JSLS*. 2014;18.

Available online at www.sciencedirect.com
ScienceDirect
journal homepage: www.JournalofSurgicalResearch.com

Association for Academic Surgery

Resident performance in complex simulated urinary catheter scenarios

Jay N. Nathwani, MD,[a] Katherine E. Law, MS,[b] Rebecca D. Ray, PhD,[a] Bridget R. O'Connell Long, BS,[a] Rebekah M. Fiers,[a] Anne-Lise D. D'Angelo, MD,[a] Shannon M. DiMarco, BA,[a] and Carla M. Pugh, MD, PhD[a,*]

[a] Department of Surgery, University of Wisconsin, General Surgery, Madison, Wisconsin
[b] University of Wisconsin, Industrial and Systems Engineering, Madison, Wisconsin

ARTICLE INFO

Article history:
Received 5 February 2016
Received in revised form
25 April 2016
Accepted 7 June 2016
Available online 16 June 2016

Keywords:
Surgical education
Urinary catheterization
Surgical assessment

ABSTRACT

Background: Urinary catheter insertion is a common procedure performed in hospitals. Improper catheterization can lead to unnecessary catheter-associated urinary tract infections and urethral trauma, increasing patient morbidity. To prevent such complications, guidelines were created on how to insert and troubleshoot urinary catheters. As nurses have an increasing responsibility for catheter placement, resident responsibility has shifted to more complex scenarios. This study examines the clinical decision-making skills of surgical residents during simulated urinary catheter scenarios. We hypothesize that during urinary catheterization, residents will make inconsistent decisions relating to catheter choices and clinical presentations.

Methods: Forty-five general surgery residents (postgraduate year 2-4) in Midwest training programs were presented with three of four urinary catheter scenarios of varying difficulty. Residents were allowed 15 min to complete the scenarios with five different urinary catheter choices. A chi-square test was performed to examine the relation between initial and subsequent catheter choices and to evaluate for consistency of decision-making for each scenario.

Results: Eighty-two percent of residents performed scenario A; 49% performed scenario B; 64% performed scenario C, and 82% performed scenario D. For initial attempt for scenario A-C, the 16 French Foley catheter was the most common choice (38%, 54%, 50%, P's < 0.001), whereas for scenario D, the 16 French Coude was the most common choice (37%, $P < 0.01$). Residents were most likely to be successful in achieving urine output in the initial catheterization attempt ($P < 0.001$). Chi-square analyses showed no relationship between residents' first and subsequent catheter choices for each scenario (P's > 0.05).

Conclusions: Evaluation of clinical decision-making shows that initial catheter choice may have been deliberate based on patient background, as evidenced by the most popular choice in scenario D. Analyses of subsequent choices in each of the catheterization models reveal inconsistency. These findings suggest a possible lack of competence or training in clinical decision-making with regard to urinary catheter choices in residents.

* *Corresponding author.* Department of Surgery, University of Wisconsin Hospital and Clinics, 600 Highland Ave, WiSOR K6/135, Madison, WI 53792. Tel.: +1 608 263-5218; fax: +1 608 253-2354.
E-mail address: pugh@surgery.wisc.edu (C.M. Pugh).

0022-4804/$ – see front matter
http://dx.doi.org/10.1016/j.jss.2016.06.021

Introduction

Every one in four hospitalized patients require a urinary catheter. The most common hospital-acquired infection comes from poor urinary catheterization insertion technique or catheter management.[1] Increased focus on the education of urinary catheterization aims to significantly decrease morbidity and health care costs associated with these hospital-acquired infections.[2-4] Hospitals have adopted stricter indications, reminder systems to remove catheters, standardized algorithms for catheter insertion, and alternative, less-invasive urinary catheter types.[5,6] These evidence-based guidelines result in decreased urinary tract infections. Nursing, residents, and staff physicians are instructed to follow these guidelines while inserting and managing urinary catheters.[5,6]

Resident adoption of the guidelines is yet unclear. With increased emphasis on patient safety and the thoughtful public concern of medical errors, surgical education is placed under the scrutiny of excellence and expectations for competency.[7] In response, surgical education incorporates simulation to provide trainees a safe, patient-free environment to develop procedure-related skills and minimize errors. In addition, in an effort to standardize education and ensure competency, the American College of Surgeons and Association of Program Directors in Surgery (ACS/APDS) created the Resident Skills Curriculum.[8] This curriculum includes a wide variety of basic and advanced skills that are considered essential for the surgical trainee to learn. In this curriculum, urinary catheterization is identified for surgical residents as a phase 1, basic/core skill, and task.

While the curriculum sets the expectation that surgical residents achieve competency in urinary catheterization, it appears that surgical residents often defer this task to a urological service when asked to troubleshoot.[9] Catheter insertion is often designated as a nursing task, and residents are often not called until a nurse fails a difficult catheterization scenario. Often when called to attempt urinary catheterization, physicians will prematurely consult a urology service with no attempt to place a catheter themselves.[9] Often times, urology services, although willing to catheterize, view some of these consults as unnecessary.[9] The suggested algorithm developed for the ACS/APDS Resident Skills Curriculum states that for an uncomplicated male or female, a 16 French (Fr) Foley catheter is an appropriate first choice.[10] This curriculum then moves forward to provide some suggestions to prevent common errors and maximize success. With premature consultation to urological services for urinary catheter insertion, it is uncertain if surgical residents have achieved competence.

The aim of this study is to assess the surgical trainee's ability to insert and troubleshoot difficult urinary catheterization scenarios in the setting of common and complex urinary pathology. We hypothesize that during urinary catheterization, residents will make inconsistent decisions relating to catheter choices and clinical presentations.

Materials and methods

Setting and participants

This study took place across seven Midwest general surgery training programs located at tertiary care hospitals. Recruitment focused on surgical residents who were entering their first year of dedicated laboratory research. Clinical residents and research residents in their second year who wished to participate were also allowed to participate. Resident participation was completely voluntary. Data collection occurred in either simulation centers or rented halls of the respective hospital's surgical department. Data were collected from residents over a period of 4 mo. The University of Wisconsin Hospitals and Clinics (UWHC) Institutional Review Board approved this study.

Research protocol

Residents filled out an initial survey with basic demographic information. After completion, they were directed to the urinary catheter station where a researcher read a standardized introduction to the urinary catheter task. Participants were encouraged to perform the urinary catheterization as they would with a real patient and were told to verbalize any actions or choices that they would normally perform but could not due to the limitations of the simulation. Residents were presented with a clinical scenario and asked to catheterize three of four simulated urinary catheter models in less than 15 min.

Each clinical scenario was common and of varying difficulty: (A) female trauma patient with a bladder injury who had bloody urine output with successful urinary catheter insertion, (B) a preoperative female who had labial constriction, (C) an elderly male who had complete urinary tract obstruction of unknown etiology, and (D) a male patient with partial blockage of the urinary tract secondary to benign prostatic hypertrophy (BPH).

Residents were provided with a full urinary catheterization kit and were told to assume the patient was appropriately prepped and draped. They were allowed five different urinary catheters to catheterize the models, including a 16 French (Fr) Foley, 16 Fr Coude, 16 Fr three-way, 14 Fr Foley, or 12 Fr Foley catheter. There was no limit to the number of catheter attempts or catheter choices participants were allowed.

Study data were collected and managed using REDCap[11] electronic data capture tools hosted at the University of Wisconsin-Madison, School of Medicine and Public Health. REDCap is a secure, web-based application designed to support data capture for research studies. Data collection focused on the catheter choices, number of catheters used, total attempts at catheterization, attempt number that led to successful urine return.

Data analysis

Data analyses were performed in SPSS.[12] Descriptive statistics and chi-square analyses were performed to explore the

relationship between initial and subsequent catheter choices and to evaluate the consistency of participants' decision-making for each scenario. A logistic regression was performed to explore the relationship between postgraduate year and success of urine return on scenario A.

Results

A total of forty-five general surgery residents (55.6% female) between postgraduate year (PGY) 2 and 4 (M = 2.7, standard deviation = 0.9) participated in the urinary catheter simulation. Residents varied in the stage of training they were in: 44.4% of individuals were entering their first research year, 22.2% were entering their second research year, and 33.4% were in their clinical year. Eighty-two percent of participants performed the scenario with female trauma, 49% performed the scenario with the preoperative female, 64% performed the scenario with the male with complete blockage, and 82% performed the scenario with the male with BPH.

The 16 Fr Foley catheter was the most popular initial choice for scenario A, B, and C at 38%, 54%, and 52%, respectively ($X^2_{(4)} = 13.95$, $P = 0.007$; $X^2_{(3)} = 16.18$, $P = 0.001$; and $X^2_{(3)} = 11.69$, $P = 0.009$). In contrast, the most popular initial choice for scenario D was the 16 Fr Coude at 37% ($X^2_{(4)} = 16.92$, $P = 0.002$). The analyses did not reveal a pattern for subsequent catheter choices after initial catheter failure as one might expect if participants were using a decision tree (Fig. 1).

Success rates associated with scenarios

Participants had the highest success rate of bladder catheterization with the preoperative female at 82%. Participants were then most successful in catheterizing the male with BPH at 40%, followed by the female trauma patient at 33%. The male with the complete blockage had no participants with successful bladder catheterization (Fig. 2). Successful urine return on the female trauma patient showed a trend of being associated with resident PGY ($X^2 = 3.75$, $P = 0.053$).

For all scenarios where bladder catheterization was achievable, participants were most likely to achieve successful bladder catheterization on their first attempt with their first urinary catheter choice. However, a few residents attempted as many as four catheters and up to five attempts with a single catheter. For the preoperative female, over 60% of participants successfully catheterized the model on their first attempt with their first catheter choice ($X^2_{(2)} = 43.03$, $P < 0.0001$). For participants who failed and made a second attempt with the same catheter, the success was approximately 10%. Even fewer participants had success when switching to the second catheter choice at 5% or less (Fig. 2).

Discussion

As urinary catheterization responsibilities have shifted to nursing staff, resident physicians are often called only in the case of complex catheterizations. This study examined resident clinical decision-making during complex urinary catheterization scenarios. We hypothesize that during urinary catheterization, residents will make inconsistent decisions relating to catheter choices and clinical presentations. Our analysis shows that first attempts at catheterization have consistent catheter choices; however, the second and third attempts have inconsistent urinary catheter choices. This suggests that for troubleshooting difficult urinary catheter scenarios, there was not a general logical decision tree among participants. For all the models where catheterization is

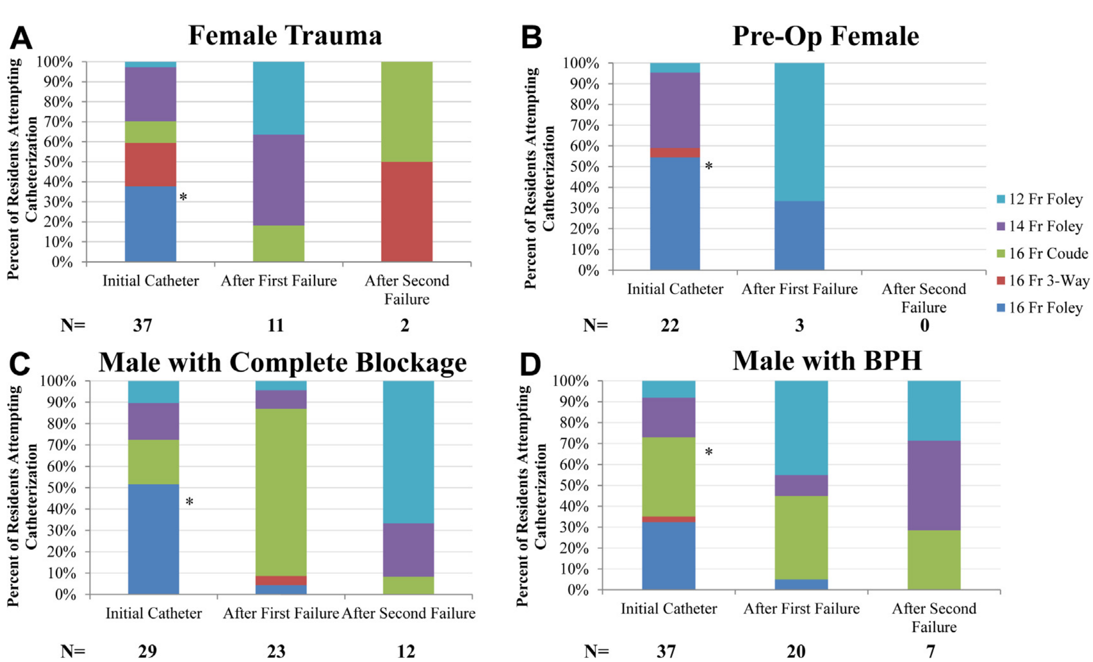


Fig. 1 – Most popular initial catheter choice followed by second and third most popular catheter choices when participants fail with initial catheter (*P < 0.01). (Color version of figure is available online.)

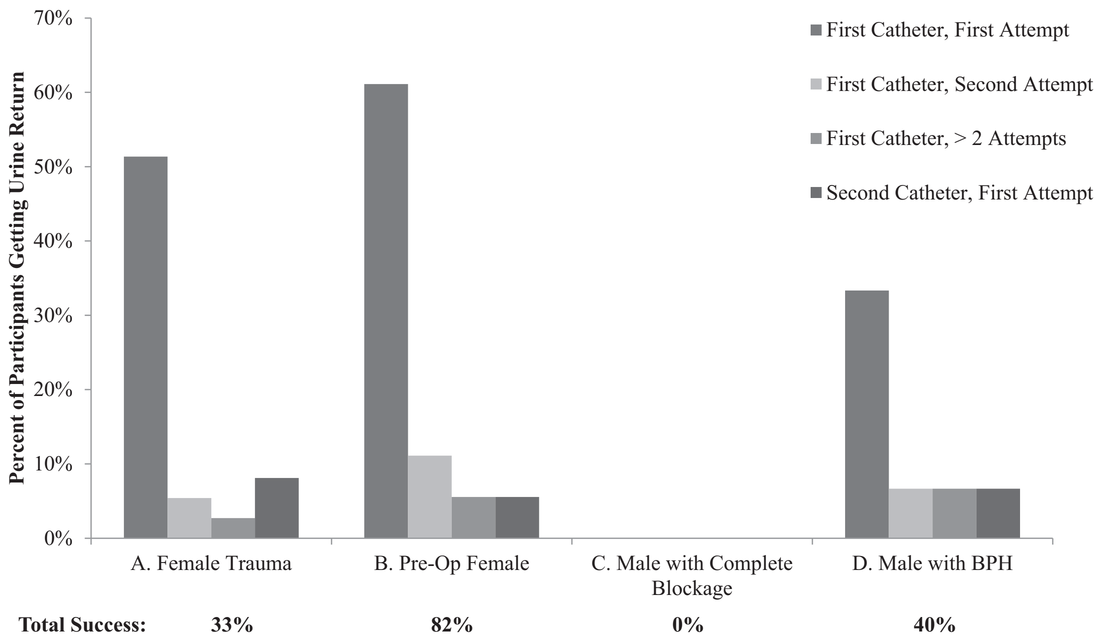


Fig. 2 – Success rates for scenario performances.

possible, we expected participants to initially try with a 14 or 16 Fr Foley catheter in the female models and a 16 Fr Foley in the male model. With initial failures, we then expected upsizing of the catheter, or, request of a Coude. Finally, if the largest catheter offered was used, we hoped that participants would verbalize upsizing to an 18 Fr Foley or Coude.[10,13] These findings may be the product of limited resident skill and education with urinary catheters.

Evaluating resident skill

Our analyses do indicate a significant pattern in initial urinary catheter choices. In scenarios A-C, residents favored the 16 Fr Foley catheter first, whereas in scenario D, residents showed a significant preference for the 16 Fr Coude catheter. The size consistency across all urinary catheter scenarios suggests that this is a popular and widely accepted initial catheter size. Choosing a 16 Fr consistently may be the product of previous teaching while in medical school or standardized algorithms adopted by hospitals. We additionally determined that participants tailored their initial catheter type according to the anticipated pathology of a given model. In particular, residents most often chose the Coude catheter first with the patient with known BPH, the only patient with a known urinary tract obstruction. This suggests that residents do strategize their catheter choice given anticipated difficulties and maximize success.

Our data show that the first catheterization attempt is most likely to predict successful urine return of all the models that were capable. For scenarios A and B, over 50% of participants were successful on the first attempt of catheterization, whereas follow-up attempts with the same catheter showed less success. When participants moved to another catheter in these scenarios, urine return success was still 10% or less. In scenario D, although the initial success rates were not as high, we still saw the general trend that the first attempt was most successful and later attempts with either the same or different catheters being less successful. This observed trend may be a reflection of participant experience. The trend of successful urine return being associated with higher PGY for scenario A supports this hypothesis. More experienced individuals may have a more successful insertion technique compared with those who had multiple attempts with multiple catheters.

The scenario success rates came out differently than expected. Specifically, we did not hypothesize that the preoperative female would have the highest success in urine return. Instead, we expected the female trauma patient to have the highest success rate, followed by the preoperative female, then male with BPH, and finally the male with complete blockage. In the actual design of our study, our preoperative female was modified to enhance the stated pathology of labial constriction, whereas the female with pelvic trauma had no physical modifications. One possibility is that each model came from different manufacturers. There may be inherent differences in the internal design of either product leading to a higher chance of success while catheterizing.

Implications for resident training

With increased concerns for patient safety, significant changes have occurred to ensure resident competency in training. For example, in 2007, the ACS/APDS Surgical Skills Curriculum was released and identified urinary catheterization as a phase 1 skill for surgical trainees, expecting that surgical trainees would learn and master this task early in training.[10] Despite the expectation of this curriculum, it appears that our participants struggle in performance of this task.

Simulation can be used to improve resident performance in urinary catheterization in a risk-free environment.

Residents are allowed infinite chances to practice their technique with no injury to a patient in a self-teaching fashion. Simulation continues to be adopted as a means of teaching in surgical education. Several validation studies show that simulation-based training can lead to improved performance and decreased patient morbidity.[14-16] The implications of a decreased complication rate could translate into decreased health care spending. A study conducted at a tertiary care facility evaluated the rates of central line–associated blood stream infections after implementing a central line insertion simulation course.[17] A significant number of central line–associated blood stream infections were prevented, resulting in a net annual savings of over seven hundred thousand dollars, resulting in a 7:1 return when considering the costs of the intervention.[17] Considering the leading cause of hospital-acquired infection currently in the United States is the urinary tract infection, it follows that proper training with urinary catheterization simulation could translate into significant health care savings by decreasing infection rates and injuries resulting from improper technique.

Interestingly, efforts to improve physician education of urinary catheter insertion have increased due to the burden that is placed on urological services for difficult catheterization scenarios. Although urinary catheterization is considered a fundamental skill and often performed by nursing staff, it appears that the majority physicians consult a urologist for difficult scenarios before attempting to place a urinary catheter themselves. In a single institutional study following urology consultation for urinary catheterization, 72% of physicians had consulted a urologist without attempting to place a catheter after a nurse failed. Interestingly, 66% of the consulted cases were successful by standard catheterization.[9] It appears that the reason for premature urology consult is multifactorial, including limited experience and poor technique.[18] In addition, as our study suggests, residents may have difficulty troubleshooting catheters after initial failure, resulting with a premature consultation to urology. To prevent unnecessary consultation, a standardized algorithm could be adopted to guide primary service teams to attempt catheterization in straightforward cases, leading to improved efficiency and primary provider performance.[9]

Although some efforts have been implemented to guide residents and decrease the burden on consulting services for urinary catheterizations, these fail to address the underlying issue of poor technique and limited experience. Moreover, current curricula do not assess resident's skill in addressing difficult scenarios in catheter placement. With the establishment of the ACS/APDS Surgical Skills Curriculum, a reasonable program for urinary catheterization exists for the development, improvement, and maintenance of catheterization skills. This preexisting curriculum could easily be expanded to include complex and difficult scenarios. Training centers could use this curriculum to ensure competency before any patient contact, early in training at the internship level. The implications of this type of intervention could be studied to better understand the full effects on patients, trainees, and health care outcomes. Of note, other well-performed studies already show the significant benefit of using simulation to improve performance and patient safety.[14-16]

Although successful in addressing the aims of the study, there were some limitations. Consulting urologists have suggested that our catheter options were limited, and that large-sized catheters should have been made available due to current standard in practice. As we previously stated, residents were encouraged to vocalize if they wanted to do something but could not. Only one resident asked for a larger size catheter. No other residents suggested that they would have used a larger catheter. In addition, while our analyses do look at success rates and their association to catheter choice, we are thoughtful of the fact that success also depends on technical skill. Future work may better test resident knowledge by allowing a wider range of urinary catheters that would be provided by verbal request. Doing so would provide further insight on the depth of surgical resident knowledge without the potential guidance of visual cues. Additional work should also focus on technical skill assessment as this may be directly related to participant success rate.

Conclusion

This study suggests that surgical residents have limited understanding on how to troubleshoot a failed urinary catheterization. It appears that first attempts at catheterization may be well understood, but, troubleshooting in the setting of failure is a skill yet to be developed. Use of simulation and development of standardized algorithms to help develop and maintain resident skill could provide guidance in straightforward scenarios and increase the chances of success.

Acknowledgment

Funding for this study came from two grants, one from the Department of Defense (W81XWH-13-1-0080) and one from a National Institute of Health grant (T32CA090217). Neither funding source had involvement in study design, data collection, analysis or interpretation, drafting, or decision to submit the article for publication.

Authors' contributions: J.N.N. was involved in data acquisition and analysis, article drafting and critical revision of the article. R.D.R. and K.E.L. conducted data analysis and performed critical revision of the article. A.-L.D.D'A., B.R.O'C.L., R.M.F., and S.M.D. were involved in data acquisition and critical revision of the article. C.M.P. performed duties of a senior author and was involved in all aspects of the study and article development.

Disclosures

The authors have nothing to disclose.

REFERENCES

1. Klevens RM, Edwards JR, Richards CL, et al. Estimating health care-associated infections and deaths in U.S. hospitals, 2002. *Public Health Rep*. 2007;122:160–166.

2. Burke J. Infection control—a problem for patient safety. *N Engl J Med*. 2003;348:651–656.
3. Saint S. Clinical and economic consequences of nosocomial catheter-related bacteriuria. *Am J Infect Control*. 2000;28:68–75.
4. Saint S, Wiese J, Amory JK, et al. Are physicians aware of which of their patients have indwelling urinary catheters? *Am J Med*. 2000;109:476–480.
5. Saint S, Kowalski CP, Kaufman SR, et al. Preventing hospital-acquired urinary tract infection in the United States: a national study. *Clin Infect Dis*. 2008;46:243–250.
6. Carter NM, Reitmeier L, Goodloe LR. An evidence-based approach to the prevention of catheter-associated urinary tract infections. *Urol Nurs*. 2014;34:238–245.
7. Scott DJ. Patient safety, competency, and the future of surgical simulation. *Simul Healthc*. 2006;1:164–170.
8. Scott DJ, Dunnington GL. The new ACS/APDS skills curriculum: moving the learning curve out of the operating room. *J Gastrointest Surg*. 2008;12:213–221.
9. Liu JJ, Guo DP, Gill H. Patterns of urinary catheter consults in a tertiary care hospital. *Can J Urol*. 2013;29:7046–7049.
10. Radomski S. (published 07/2007) ACS/APDS Surgical Skills Curriculum For Residents, Phase 1: Module 6 Urethral and Suprapubic Catheterization. Available at: https://www.facs.org/education/program/apds-resident. Accessed February 12, 2015.
11. Harris PA, Taylor R, Thielke R, Payne K, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377–381.
12. BM Corp. Released 2013. IBM SPSS statistics for Windows, version 22.0. Armonk, NY: IBM Corp.
13. Ghaffary C, Yohannes A, Villanueava C, Leslie SW. A practical approach to difficult urinary catheterizations. *Curr Urol Rep*. 2013;14:565–579.
14. Scott DJ, Bergen PC, Euhus DM, et al. Laparoscopic training on bench models: better and more cost effective than operating room experience? *J Am Coll Surg*. 2000;191:272–283.
15. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves OR performance: results of a randomized, double-blinded study. *Ann Surg*. 2002;236:458–463.
16. Fried GM, Feldman LS, Vassiliou MC, et al. Proving the value of simulation in laparoscopic surgery. *Ann Surg*. 2004;240:518–525.
17. Cohen ER, Feinglass J, Barsuk JH, et al. Cost savings from reduced catheter-related blood stream infection after simulation based educations for resident in a medical intensive care unit. *Simul Healthc*. 2010;5:98–102.
18. Thomas A, Giri S, Meagher D, Creagh T. Avoidable iatrogenic complications of urethral catheterization and inadequate intern training in a tertiary-care teaching hospital. *BJU Int*. 2009;104:1109–1112.

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.JournalofSurgicalResearch.com

Association for Academic Surgery

Errors in bladder catheterization: are residents ready for complex scenarios?

Bridget R. O'Connell-Long, BS, Rebecca D. Ray, PhD, Jay N. Nathwani, MD, Rebekah M. Fiers, BS, and Carla M. Pugh, MD, PhD*

Department of Surgery, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, Wisconsin

ARTICLE INFO

Article history:
Received 4 February 2016
Received in revised form
21 April 2016
Accepted 2 June 2016
Available online 9 June 2016

Keywords:
Simulation
Performance
Confidence
Surgery
Education

ABSTRACT

Background: The aim of this study was to investigate whether junior surgical residents had successfully mastered bladder catheterization. Our hypothesis was that surgical residents would be overly confident in their abilities and underestimate the potential for case complexity.

Materials and methods: PGY 2-4 surgery residents ($n = 44$) were given 15 min. to complete three of four bladder catheterization simulations. Participants reported their mastery by rating confidence using a 5-point Likert scale. Multiple linear regression analysis was used to test predictors of procedure performance.

Results: Participants made a total of 228 errors with an average of 5.1 errors (standard deviation = 2.6) per participant. The most common errors included not maintaining the sterile field (52.0%), failure to get urine return (20.3%), and inflating the catheter balloon before urine return (8.4%). Some residents committed the same error more than once. Presimulation confidence ratings ranged from "1" being not confident to "5" being extremely confident. Average presimulation confidence was 4.42 (range 1-5, standard deviation = 0.85). Sixteen (36%) residents ranked their presimulation confidence in problem-solving abilities as "moderately confident" or below, whereas 28 (64%) were "very confident" or above. The lower the resident's presimulation confidence in problem-solving, the more errors they committed during the simulation (beta = −0.33, t = −2.15, $P = 0.04$).

Conclusions: The residents did not perform as well as they anticipated when presented with more complicated bladder catheterization scenarios. Simulation can be used to identify and expose potential errors that may occur during complex presentations of basic procedures. This type of training and assessment may facilitate mastery.

Introduction

Catheter-associated urinary tract infections (CAUTI) are currently the number one hospital-acquired infection in the United States. The CDC estimates that CAUTI attributes to 13,000 deaths and at least $400 million in additional cost per year nationally. The detrimental effects of CAUTI have caused the CDC to define a set of Core Prevention strategies with high-level evidence for prevention. Proper training of personnel and maintaining sterile insertion technique are included in

* *Corresponding author*. Department of Surgery, University of Wisconsin Hospital and Clinics, 600 Highland Ave, K6/135 CSC, Madison, WI 53792. Tel.: +1 262 470 7855; fax: +1 262 642 5146.
E-mail address: boconnelllon@wisc.edu (B.R. O'Connell-Long).
0022-4804/$ – see front matter
http://dx.doi.org/10.1016/j.jss.2016.06.005

the Core Prevention strategies as defined by the Centers for Disease Control and Prevention.[1-3] Although nurses perform a majority of urinary bladder catheterizations in the hospital, physicians are often called on for more complex placements. The Association for Healthcare Research and Quality has released a safety program dedicated to the prevention of device-associated infections. The safety program focuses on training resident physicians to be CAUTI-prevention experts, as they are often the first physicians to see the patient.[4]

The current body of literature indicates that simulation training can enhance residents' confidence and performance in performing surgical and bedside procedures.[5-7] However, simulation programs can vary widely among medical schools and residency training programs and do not always include complex presentations of procedures for training and assessment.[8,9] In 2007, the ACS/APDS Surgical Resident Skills Curriculum was developed in an attempt to standardize training of surgical residents. According to the ACS/APDS curriculum, urethral catheterization should be mastered during the PGY1 and 2 y.[10]

The aim of this study was to investigate whether surgical residents had successfully mastered bladder catheterization. We intended to determine the level of mastery of this basic procedure by presenting more complex and well-known clinical scenarios. Our hypothesis was that surgical residents would be overly confident in their abilities and underestimate the potential for case complexity.

Materials and methods

Setting and participants

General surgery residents (PGY 2-4, $n = 44$, 55.6% female, 44.4% male) from seven Midwest training programs participated in this study. Surgical residents in their dedicated research years made up 60% of the participants. Study participants were given 15 min to complete three bladder catheterization simulations. Data were captured using motion-tracking software and video and audio recording. Participants were double gloved using their normal size surgical gloves. Motion-tracking wires were secured on their hands between the two gloves. Participants were given no feedback on their performance after completion of the simulation. This study was approved by the University of Wisconsin-Madison Social and Behavioral Sciences Institutional Review Board.

Bladder catheterization simulations

Study participants completed three of four randomized bladder catheterization simulations as part of a larger study. The simulations represented well-known clinical scenarios: a female trauma patient, a female preoperative patient, a preoperative male with a full urethral blockage, and a male with urinary retention due to benign prostatic hypertrophy (Table 1). This study used modified Limbs & Things and Nasco bladder catheterization models. The female trauma patient was a standard model that returned bloody urine to indicate prior trauma. The female preoperative patient simulation was created by adding sutures to the inside of labia to simulate a labial stricture. The penile urethra was completely tied off inside the model to simulate a male with complete urethral blockage. Finally, the penile urethra in male with partial urethral structure was partially obstructed using a polyethylene tube.

Table 1 – Description of the four simulated bladder catheterization procedures.

Description	Clinical scenario	Unknown pathology
Female trauma	Motor vehicle accident, pelvic fracture	Bladder injury
Preoperative female	None	Labial constriction
Preoperative male	Rectal cancer, preoperative lower anterior resection	Complete obstruction
Male, retention	Benign prostatic hypertrophy	None

Participants were given a short description of the clinical scenario that corresponded with each simulation. Participants were provided with a simulated catheterization kit including sterile water, lubricant, and a choice of five different catheter sizes. The residents were told to assume that the field was sterile, and the patient had been prepped. In addition, participants were given the option to place a urology consult at any point in the simulation.

Participant data

One investigator reviewed video and audio data from each participant and evaluated procedure performance. Errors were defined using the ACS/APDS Technical Skills Curriculum.[10] Investigators used a standard rubric from the ACS/APDS Technical Skills Curriculum to evaluate errors. Error definitions are shown in Figure 3. The residents were asked to complete a survey both before and after simulation. They rated the perceived procedure difficulty and their personal confidence level using a 5-point Likert scale for three categories: identifying relevant anatomy, problem-solving, and

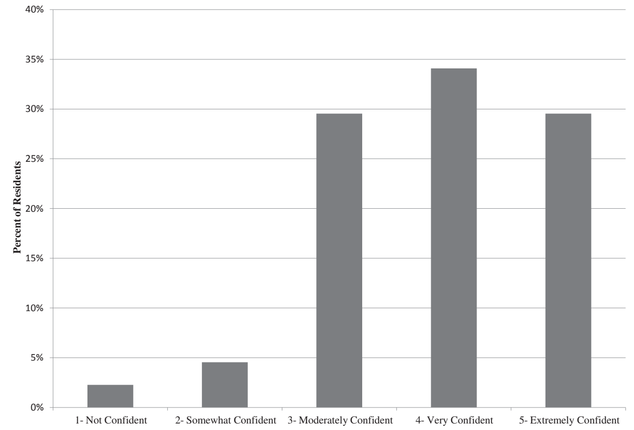


Fig. 1 – Resident presimulation confidence in problem-solving ability for urinary catheterization.

completing the entire surgical task (Fig. 1). They also rated the amount of skill reduction they expected as a function of their dedicated time in laboratory research.

Data analysis

Study data were collected and managed using REDCap electronic data capture tools hosted at the University of Wisconsin-Madison, School of Medicine and Public Health.[11] All analyses were performed using SPSS.[12] Multiple linear regression analysis was used to model perceived mastery as predictors of procedure performance as measured by errors and time to call for a urology consult.

Results

Presimulation confidence

Presimulation confidence ratings ranged from "1" being not confident to "5" being extremely confident. Average presimulation confidence was 4.42 (range 1 to 5, standard deviation [SD] = 0.85). There was no difference between PGYs in self-reported confidence. The lowest PGY average was PGY2 (3.67), and the highest was PGY3 (4.83). Sixteen (36%) residents ranked their presimulation confidence in problem-solving abilities as "moderately confident" or below, whereas 28 (64%) were "very confident" or "extremely confident" (Fig. 2). The lower the resident's presimulation confidence in problem-solving, the more errors they committed during the simulation (beta = −0.33, t = −2.15, P = 0.04). Participants with higher presimulation confidence in problem-solving abilities took less time before deciding to place a urology consult (beta = −1.53, t = −4.32, P = 0.001).

Post-simulation confidence

Post-simulation confidence ratings ranged from "1" not confident to "5" extremely confident. Average post-simulation confidence was 3.56 (SD = 0.81), with 3 = moderately confident and 4 = very confident (Fig. 1). There was no relationship between post-simulation confidence level and the total errors ($P > 0.05$).

Errors

Participants made a total of 228 errors with an average of 5.1 errors (SD = 2.6) per participant. The average number of errors per participant was 1.95 for female trauma, 1.29 for female preop, 1.97 for the male with partial urethral blockage, and 1.73 for male with complete urethral blockage (Fig. 2). The most common errors included not maintaining the sterile field (52.0% of errors), failure to get urine return (20.3% of total errors), and inflating the catheter balloon before urine return (8.4% of total errors; Fig. 3). Residents who perceived a greater reduction in technical skills during their dedicated research years made fewer errors (beta = −0.42, t = −2.76, P = 0.01). Twenty-three (53.3%) of residents failed to inflate the catheter balloon after urine return, 18 (40.0%) failed to maintain the sterile field, and 14 (33.3 %) did not use lubricant (Fig. 4). Examples of common errors are shown in Figure 5.

Discussion

This study investigated how well surgical residents had mastered urinary bladder catheterization. We found that residents did not perform, and they anticipated when presented with more complicated bladder catheterization scenarios. We assumed that if surgical residents had reached a level of mastery, we would see almost no errors during the simulation. Our data indicate that they have not mastered this basic, high-volume procedure despite relatively high confidence levels before the simulation. In addition, the most common errors seen during the simulation (Fig. 3) have been shown to contribute directly to the development of CAUTI.[2,3]

Our study found that the most common error among residents was failure to maintain the sterile field. We found that 40% of participants committed this error, indicating that this is not a result of multiple errors by a few individuals. These

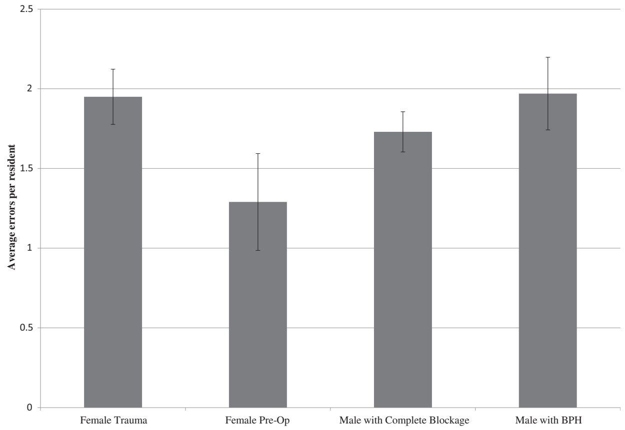


Fig. 2 – Average number of errors per resident for each bladder catheter simulator.

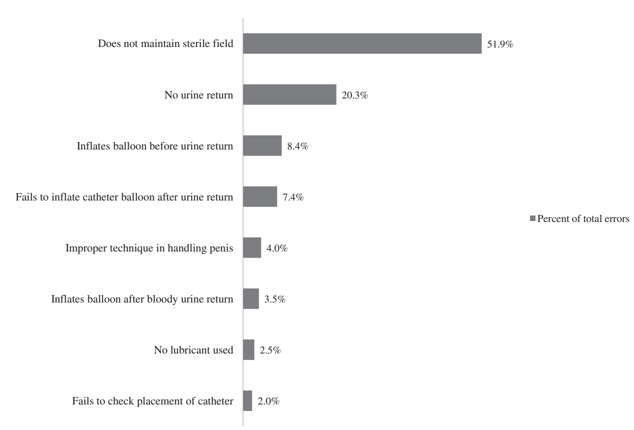


Fig. 3 – Composition of errors committed by surgical residents across all four models.

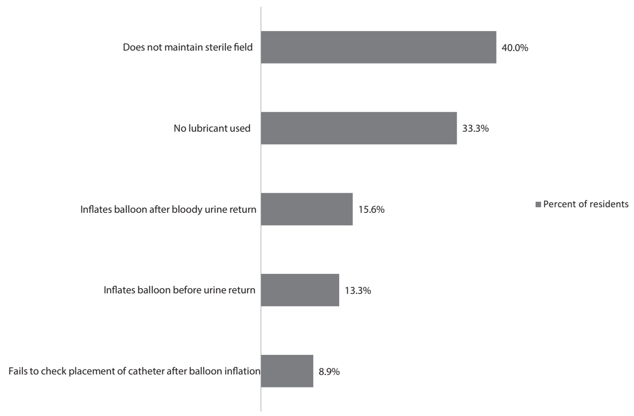


Fig. 4 – Percent of surgical residents committing the most common errors.

results are significant as maintaining sterile insertion technique shows high-level evidence for the prevention of CAUTI. Two other common errors among study participants could contribute to increased risk of infection: inflating the catheter balloon before urine return and inflating the balloon after bloody urine return (Fig. 4). Inflating the balloon before placement in the bladder can cause trauma to the urethral tissue and facilitate the spread of bacteria into the bloodstream. In addition, the bloody urine indicates previous trauma and warrants a consult and potential removal of the indwelling catheter. According to the New England Journal of Medicine guidelines, known or suspected urethral injury is an absolute contraindication to urethral catheterization.[13] Our results indicate that surgical residents may be unaware or noncompliant with these established guidelines regarding insertion of urinary catheters.

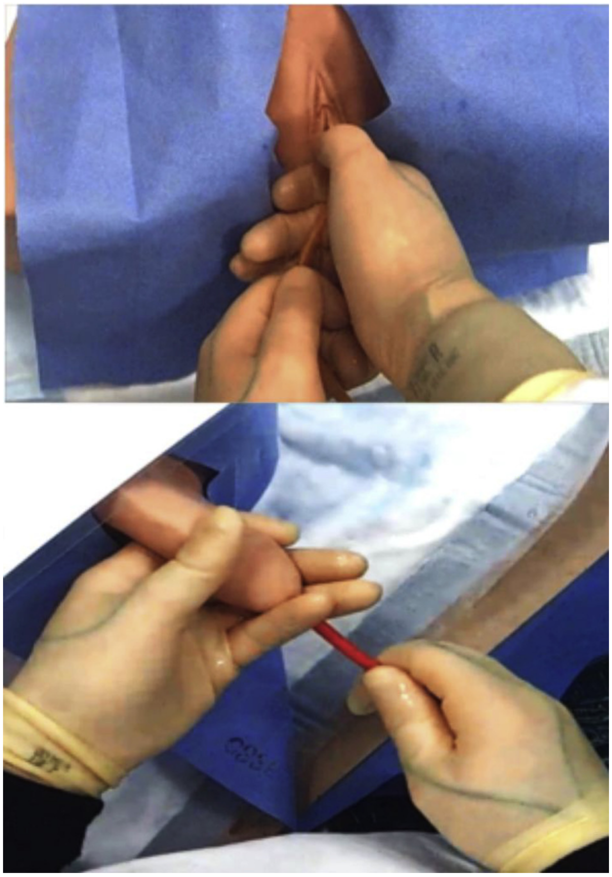


Fig. 5 – Example of errors in maintaining the sterile field. (Color version of figure is available online.)

This study is limited by our ability to use simulation to measure clinical performance, as participants' behavior during a simulated procedure may differ from the clinical setting. While participants were instructed to treat the simulation as a real-life patient and environment, performance anxiety, video recording and motion-tracking equipment, competition among residents, and the presence of the researchers could have influenced the behavior of our study participants. These factors may limit our ability to generalize this study to real-life performance. Due to this, it is difficult to determine if the most common errors among residents in our study are an accurate reflection of those made in a clinical setting. However, many clinical and academic settings have similar distractions, and we still expect surgical residents to have reached a level of automaticity for this procedure where they would be unaffected by these distractions.

One potential barrier to mastery is the amount of repetition that junior residents are exposed to during their training. Because urinary catheterization has primarily become a nursing procedure, residents may not perform the volume of procedures that are needed to adequately assess and problem solve more complex presentations.[14] Our results show that 36% of junior residents reported "moderate confidence" or below with respect to problem-solving for urinary catheterization. As expected, we found that lower presimulation confidence in problem-solving resulted in a higher number of errors during the simulation. However, our study also found that residents who perceived a greater reduction in technical surgical skills during their dedicated research years made fewer errors during the bladder catheterization simulation. This may mean they were more careful during the simulation as a result of believing that they had lost some technical skills. These findings support the need to develop standardized and concrete skills assessments to ensure that all junior residents have mastered their skills curriculum. Using simulation to assess resident performance could help reduce patient morbidity and cost related to unnecessary errors.

Conclusions

The results of this study demonstrate a need to develop a culture in surgery where residents are able to address gaps in knowledge and practice basic procedures until they reach a level of mastery. Simulation is a valuable tool that allows us to identify areas for improvement without the risks to patients. In addition, simulation can be used to help residents identify and expose potential errors that may occur during complex presentations of basic procedures. This type of training and assessment is necessary and can facilitate mastery.

Acknowledgment

The author would like to acknowledge Shannon DiMarco for contributions to data collection and critical revision of the manuscript.

This study received funding from the National Institutes of Health grant T35DK062709 and the Department of Defense grant W81XWH-13-1-0080. Neither funding source had involvement in study conception or design, data analysis, writing of the article, or decision to submit the article to publication.

Authors' contributions: BO'C-L performed data analysis and interpretation, writing the article, and critical revision of the article with significant contribution and assistance from RR. RF contributed to data collection, data analysis and interpretation, and critical revision of the article. JN contributed to data analysis and interpretation and critical revision of the article. CMP performed duties of a senior author; contributed to study conception and design; data collection, analysis and interpretation; writing the article, and critical revision of the article.

Disclosure

The authors reported no proprietary or commercial interest in any product mentioned or concept discussed in the article.

REFERENCES

1. Chenoweth C, Saint S. Preventing catheter-associated urinary tract infections in the intensive care unit. *Crit Care Clin*. 2013;29:19–32.
2. Gould CV, Umscheid CA, Agarwal RK, Kuntz G, Pegues DA, Patrick J. Guideline for prevention of catheter-associated urinary tract infections 2009. *Infect Control Hosp Epidemiol*. 2009;31:319–326.
3. Cochran S. Care of the indwelling urinary catheter: is it evidence based? *J Wound Ostomy Continence Nurs*. 2007;34:282–288.
4. for Healthcare Research A. AHRQ Safety Program for Reducing CAUTI in Hospitals Implementation Guide 2 n.d.
5. Cooke JM, Larsen J, Hamstra SJ, Andreatta PB. Simulation enhances resident confidence in critical care and procedural skills. *Fam Med*. 2008;40:165–167.
6. Smith CC, Huang GC, Newman LR, et al. Simulation training and its effect on long-term resident performance in central venous catheterization. *Simul Healthc*. 2010;5:146–151.
7. Sturm LP, Windsor JA, Cosman PH, Cregan P, Hewett PJ, Maddern GJ. A systematic review of skills transfer after surgical simulation training. *Ann Surg*. 2008;248:166–179.
8. Scott DJ, Dunnington GL. The new ACS/APDS skills curriculum: moving the learning curve out of the operating room. *J Gastrointest Surg*. 2008;12:213–221.
9. Bréaud J, Chevallier D, Benizri E, et al. The place of simulation in the surgical resident curriculum. The pedagogic program of the Nice Medical School Simulation Center. *J Visc Surg*. 2012;149:e52–e60.
10. ACS/APDS Surgery Resident Skills Curriculum n.d. Available at: https://www.facs.org/education/program/apds-resident. Accessed January 4, 2016.
11. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377–381.
12. IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.
13. Male Urethral Catheterization — NEJM n.d. Available at: http://www.nejm.org/doi/full/10.1056/NEJMvcm054648. Accessed January 25, 2016.
14. Kolozsvari NO, Feldman LS, Vassiliou MC, Demyttenaere S, Hoover ML. Sim one, do one, teach one: considerations in designing training curricula for surgical simulation. *J Surg Educ*. 2011;68:421–427.

Available online at www.sciencedirect.com
ScienceDirect
journal homepage: www.JournalofSurgicalResearch.com

Association for Academic Surgery

Do resident's leadership skills relate to ratings of technical skill?

Samantha J. Gannon, BS,[a] *Katherine E. Law, MS,*[b] *Rebecca D. Ray, PhD,*[a]
Jay N. Nathwani, MD,[a] *Shannon M. DiMarco, BA,*[a]
Anne-Lise D. D'Angelo, MD,[a] *and Carla M. Pugh, MD, PhD*[a,b,*]

[a] *Department of Surgery, University of Wisconsin–Madison, Madison, Wisconsin*
[b] *Department of Industrial and Systems Engineering, University of Wisconsin–Madison, Madison, Wisconsin*

ARTICLE INFO

Article history:
Received 5 February 2016
Received in revised form
31 July 2016
Accepted 10 August 2016
Available online 19 August 2016

Keywords:
Nontechnical skill
Technical skill
Simulation
Education
Bowel repair

ABSTRACT

Background: This study sought to compare general surgery research residents' survey information regarding self-efficacy ratings to their observed performance during a simulated small bowel repair. Their observed performance ratings were based on their leadership skills in directing their assistant.

Methods: Participants were given 15 min to perform a bowel repair using bovine intestines with standardized injuries. Operative assistants were assigned to help assist with the repair. Before the procedure, participants were asked to rate their expected skills decay, task difficulty, and confidence in addressing the small bowel injury. Interactions were coded to identify the number of instructions given by the participants to the assistant during the repair. Statistical analyses assessed the relationship between the number of directional instructions and participants' perceptions self-efficacy measures. Directional instructions were defined as any dialog by the participant who guided the assistant to perform an action.

Results: Thirty-six residents (58.3% female) participated in the study. Participants who rated lower levels of decay in their intraoperative decision-making and small bowel repair skills were noted to use their assistant more by giving more instructions. Similarly, a higher number of instructions correlated with lower perceived difficulty in selecting the correct suture, suture pattern, and completing the entire surgical task.

Conclusions: General surgery research residents' intraoperative leadership skills showed significant correlations to their perceptions of skill decay and task difficulty during a bowel repair. Evaluating resident's directional instructions may provide an additional individualized intraoperative assessment metric. Further evaluation relating to operative performance outcomes is warranted.

* *Corresponding author.* Department of Surgery, University of Wisconsin Hospital and Clinics, 600 Highland Ave., K6/135 CSC, Madison, WI 53792. Tel.: +1 608 263 5218.
E-mail address: pugh@surgery.wisc.edu (C.M. Pugh).
0022-4804/$ – see front matter
http://dx.doi.org/10.1016/j.jss.2016.08.050

Introduction

Success in the operating room (OR) has most often been attributed solely to technical ability. Presently, nontechnical skills are being recognized as contributing to operative success.[1] Nontechnical skills are defined as the cognitive and interpersonal skills that reinforce a surgeon's technical ability that help contribute to safe practice and successful surgical outcomes.[2] A surgeon's skill directing others in the OR may be as critical to a patient's surgical outcome as their skills with a scalpel.[3] Emphasized training in these nontechnical skills have shown to benefit OR performance in a variety of ways.[1] The observational teamwork assessment for surgery showed that increased teamwork development lead to fewer adverse events during surgery.[4] Similar results were found in a review of communication skills and error frequency.[5] Improved communication and leadership among those on a surgical team increases effectiveness in the OR by decreasing operative time and improving patient safety.[6-9]

While it may seem that nontechnical skills are only complementary to technical skills, clinical residency training should also focus on teaching nontechnical skills because of the benefits demonstrated intraoperatively.[4-9] However, nontechnical skills can be harder to assess[10] and are commonly minimized in training.[11] In these cases, training using simulation is a feasible option to teach nontechnical skills that may lie outside the typical clinical environment[12] or are considered too risky for practice on patients.[13] Surgical simulation has also been used for assessments in decision-making,[14] error training and management,[15] measuring operative independence,[16] and resident readiness for the OR.[16] During postsimulation training, residents expressed that simulation training helped minimize skills decay and maintain confidence during time away from clinical training for dedicated research.[17] Further results showed that resident ratings of confidence and skill independence increased after simulation training.[18] More generally, Campbell *et al.*[19] found that simulation training led to an increase in both surgical performance and self-efficacy measures.

Ongoing work in our laboratory using discourse analysis suggests that communication patterns between a surgeon and operative assistant are strongly associated with technical skills. Additional results showed that discourse analysis may provide important insight on surgical ability and competence. In a simulated laparoscopic ventral hernia repair, senior surgical resident communication with their operative assistant was coded into two categories: giving instructions and asking questions.[20] Discourse analysis showed that giving the operative assistant instructions was an important element of performance and was linked to higher scores in final product analysis. Although this study controlled for individual differences in discourse, it did not evaluate participants' self-efficacy or how self-efficacy may relate to providing directional instruction to operative assistants.

In the present study, measures of participant's self-efficacy in their technical (small bowel repair) skills and nontechnical skills (operative leadership) were included. Breaking down operative leadership further, we look specifically at how these leadership skills show through as the participants direct their operative assistant during the surgical task. We believe that this indirectly relates to the level of knowledge of the trainee and their leadership skills, as "use of assistant" has been a validated topic included in the Objective Structured Assessment of Technical Skills (OSATS) assessment form. These measures give an additional layer to the objective assessment of performance and provide context regarding participant's behaviors and whether they are indicative of subjective skill perception.

The aim of this study was to understand research residents' nontechnical leadership skills as they compare to their self-perceptions of technical and nontechnical skill decay, task difficulty, and confidence in the necessary procedural steps for repairing a small bowel. Operative leadership was defined as the participant's ability to give directional instructions to an operative assistant during the procedure. We defined directional instruction to be any dialog by the participant who guided the assistant to perform an action. Our goal is to assess the utility of "giving instructions" as a useful performance metric and to understand if this performance metric has any relation to self-assessments of technical skill. We hypothesized that participants' expectations for skill decay, difficulty, and confidence in their repair would correlate to their ability to give directional instruction to the operative assistant during a simulated small bowel repair.

Methods

Population

This study included 40 (57.5% female) general surgery residents (Post Graduate Year 2-5) in dedicated research. Four participants were excluded due to incomplete data. The remaining 36 participants (58.3% female) were being trained in six programs across the Midwest. This study was part of a larger, longitudinal multiprocedure investigation aimed at identifying surgical skills that degrade over periods of nonuse. The University of Wisconsin Hospitals and Clinics Institutional Review Board approved this study.

Preparticipant survey

Participants were first given a survey to collect general demographic information on their training and experience.[16] The survey asked participants to rate their perception of surgical skillset reduction related to intraoperative decision-making and small bowel repair skills using a five-point Likert scale (1 = no reduction, 5 = very large reduction). Participants were then asked to rate their perceived difficulty (1 = not difficult, 5 = extremely difficult) in completing three procedural steps of a small bowel repair which included selecting the correct suture type, suture pattern, and successfully performing the entire surgical task. The survey also asked participants to rate their confidence (1 = not confident, 5 = extremely confident) in these same procedural skills.

Research protocol

Participants were presented with a simulated scenario of a patient with multiple gunshot wounds to the abdomen.

Participants were asked to locate, inspect, and repair the injured portion of bovine small bowel. The injury contained a small and large enterotomy to the antimesenteric portion of the small bowel. The injuries were located close in proximity but differed in size and character. The larger injury (1 × 1 cm) was simulated to have jagged edges, whereas the smaller injury (0.5 × 0.5 cm) was made to have cleaner edges. Participants were given 15 min to complete their repair. Audio and video data were recorded during each participant's performance.

Instruments

Participants were given access to all necessary surgical instruments to repair the injuries with the exception of stapling devices and electrocautery. In addition, they were given the choice of five different types of 3-0 suture. A researcher served as an operative assistant and was considered to be at the level of a medical student. This assistant was instructed to give scripted responses to clarify study expectations of the participant; however, they were restricted from providing feedback to assist with the participant's repair.

Discourse analysis

Operative leadership was assessed using the number of directional instructions the participant provided to their operative assistant during their 15-minute repair. Coding was conducted by a single, blinded independent observer and was grounded in the leadership language used by two commonly cited team surveys: (1) nontechnical skills for surgeons (NOTSS) and (2) nontechnical skills (NOTECHS).[21,22] The observer coded all the directional instructions that each participant provided to the operative assistants. "Directional instruction" represented those instances when the participant guided the operative assistant to perform a specific task or motion. These instructions generally fell into two broad categories; either direction for an instrument, or direction for motion. An example of a direction for an instrument was, "Can I get a pickup please, without teeth?" whereas an example for direction of motion was, "Alright so I want you to cut near the bowel over here but not into the bowel."

Data analysis

Using SPSS 23, correlations were performed between the number of directional instructions given to the operative assistant and participants' preprocedure ratings of skill reduction, difficulty, and confidence.

Results

Population

Participants in this study were, on average, 8.27 mo into their dedicated research time (standard deviation = 8.91) (Fig. 1). More than half of the participants (56.8%) were in their first 4 mo at the time of the study, whereas some (10.8%) had already completed more than 2 y of research.

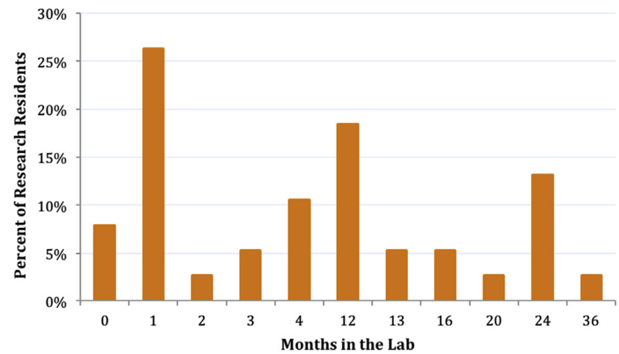


Fig. 1 – The number of mo participants had spent in dedicated research at the time of collection. (Color version of figure is available online.)

Discourse analysis

Participants gave between 3-45 directional instructions to their operative assistant throughout the 15-minute repair (Fig. 2). On average, the participants gave their operative assistants 16 directional instructions (standard deviation = 10.88). Most participants (86%) gave between 3-30 instructions.

Preparticipant survey

Perception in surgical skillset decay

Our results revealed a trend between the number of instructions and decay in decision-making skills. Specifically, those who gave a higher number of instructions anticipated less reduction in their decision-making skills [$r(38) = -0.335$, $P = 0.055$; Fig. 3]. Participants who gave a larger number of directions to their assistant perceived a lower reduction in their small bowel repair skills [$r(38) = -0.397$, $P = 0.017$; Fig. 3].

Perceived difficulty

A correlation was found between number of directional instructions given to the assistant and perceived difficulty ratings in all three components of the procedure. This relationship revealed that the participants who provided more instructions to their assistant perceived less difficulty in selecting the correct suture, selecting the correct suturing pattern, and performing

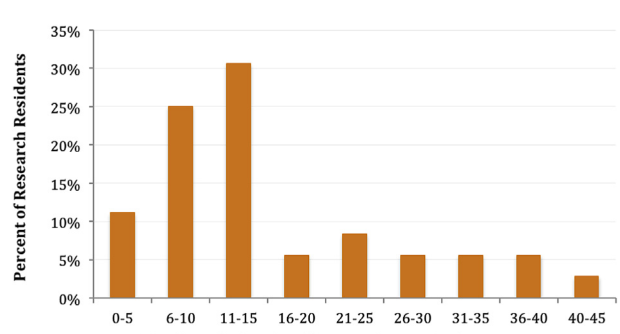


Fig. 2 – The number of directional instructions participants gave their operative assistants during the simulated small bowel repair. (Color version of figure is available online.)

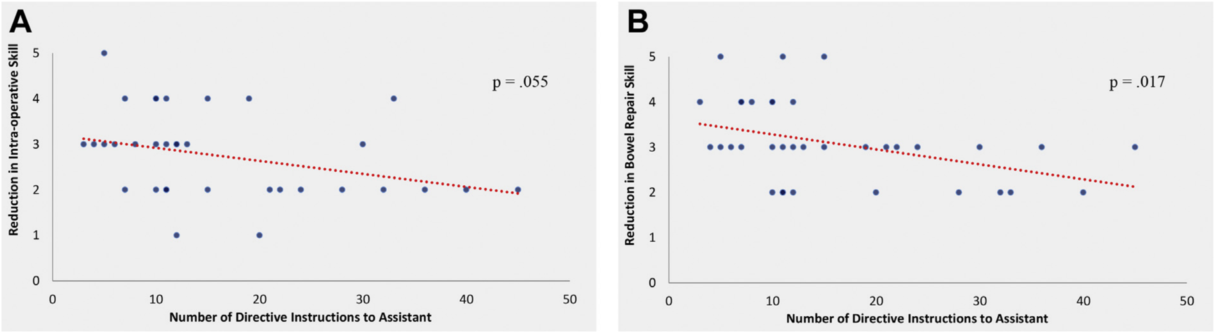


Fig. 3 – The number of instructions given by participants relating to ratings of perceived reduction in: (A) intraoperative decision-making skills identified as a trend and (B) bowel repair skills identified as a significant negative correlation. (Color version of figure is available online.)

the entire small bowel repair [suture r(35) = −0.422, P = 0.012; suturing pattern r(35) = −0.495, P = 0.002; entire surgical procedure r(35) = −0.403, P = 0.016; Fig. 4].

Confidence

There was no relationship between confidence and number of directional instructions given by the participant to the assistant during the repair [r(38) = 0.137, P = 0.426].

Discussion

Residents who were taking time away from clinical training to participate in research were given a survey to rate their self-efficacy measures regarding both technical and nontechnical skills related to a small bowel repair. This study focused on participants' operative leadership skills with respect to skill decay, task difficulty, and confidence. Operative leadership was assessed using the number of directional instructions the participant provided to their operative assistant during their 15-minute repair.

Results showed that residents who gave more direction, via a higher number of instructions, expected to have less decay in both small bowel repair and intraoperative decision-making skills. Specifically, participants who gave more instruction to their assistant tended to rate their reduction in intraoperative decision-making skills as having only mildly reduced. Likewise, those participants who were able to

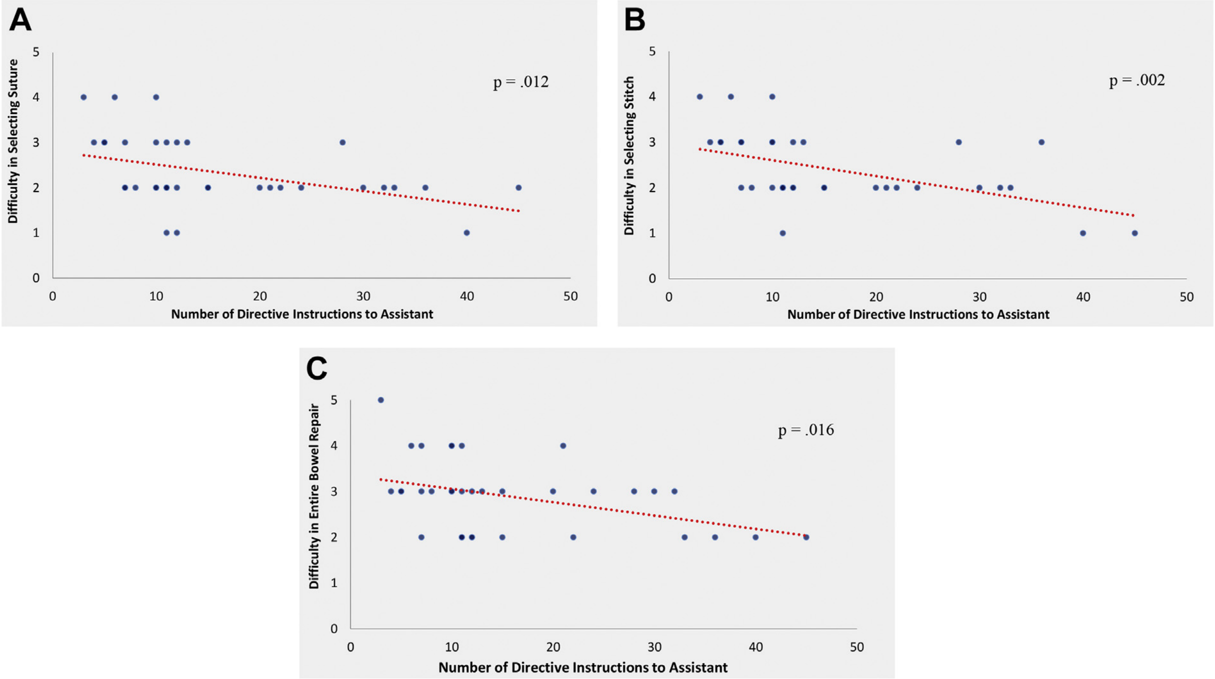


Fig. 4 – The number of instructions given by participants relating to perceived difficulty ratings in (A) selecting the correct suture type, (B) selecting the correct suturing pattern, and (C) completing the entire small bowel repair all identified as significant negative correlations. (Color version of figure is available online.)

provide more direction to their assistant were significantly less concerned with their perceptions of decay regarding their small bowel repair skills.

In procedural steps related to the bowel repair, residents who provided more direction to their operative assistant rated significantly lower difficulty in all steps related to the repair. These steps included selecting the correct suture, selecting the correct suturing pattern, and completing the entire small bowel repair. Higher difficulty ratings in these procedural steps were correlated with reduced operative leadership skills in directing their assistant to help during the bowel repair.

When assessing how participants rated their confidence in these specific technical and nontechnical skills, there was no relationship between self-efficacy ratings and the number of directional instructions given. A likely explanation for this finding is that participants did not feel a change in their confidence during their time away for research. In addition, more than half of the participants had only been away from clinical training for 4 mo or less. This discrepancy in our findings may show that blinded observer ratings of a small component of nontechnical skills such as leadership, may be a more sensitive indicator of performance than self-ratings. As this study was part of a larger longitudinal study, future analysis of participants who choose to complete the study again will afford us the ability to reassess their confidence after spending more time in the laboratory, away from clinical training.

Two of the most commonly used assessment tools for measuring nontechnical skills include NOTSS and NOTECHS.[21,22] A limitation in using these tools for our study is that we had a very scripted, limited team interaction. There was one surgeon and one assistant, and the assistant was instructed not to give guidance or suggestions to the surgeon. This limited and structured interaction as well as the time limitation of 15 min did not provide sufficient context to fully evaluate participants in the team environment. While NOTSS is best for assessing individual team members, we note that both assessments include a leadership anchor. NOTSS focuses on "supporting others" and "coping with pressure." NOTECHS focuses on how much the leader involves other team members in the course of action. Given our study constraints, we evaluated a small component of operative leadership—providing direction to an operative assistant. This was measured using the number of instructions given and then compared with participants' self-efficacy ratings. Analysis showed a correlation between the nontechnical components of intraoperative decision-making and participant's perceptions of technical skill decay and difficulty. As such, nontechnical skills may be a useful indicator of overall operative skill. In addition, trainee instruction in leadership and nontechnical skills may facilitate learning and/or execution of technical tasks.

While it remains unclear which step is more critical to develop first, studies show that leadership and communication skills may enhance other technical skills. In a study looking at nontechnical skills in the OR, Parker *et al.*[23] identified guiding and supporting as the most frequent operating leadership behavior, followed by communicating, coordinating, task management, and decision-making. Avenues that have been used to develop such nontechnical skills include increased teaching opportunities,[24] OR coaching experiences,[2] as well as training that focuses on teaching interpersonal,[25] and communication skills.[26] As evident by our results, increasing these nontechnical skills may simultaneously show benefits of increased technical skill.

In summary, the largest limitation of the present study is the lack of ability to use a well-known and validated tool for measuring nontechnical skill. Our central aim was to validate a specific component of nontechnical skill on an individual level as this was shown to be important in prior studies in our lab in a different setting.[20] A second limitation of this study was assessment of very few technical (bowel repair skills) and nontechnical skills (operative leadership and decision-making). We acknowledge that these skills represent only a small minority of the extensive skills required of general surgery residents. Further limitations included a lack of formal measurement and control of the amount of talking for each participant as this could influence the number of instructions the participant might have given their assistant.

Conclusion

Our results support the use of directional instructions as a potential metric to evaluate intraoperative leadership skills. Nontechnical skills training during residency could lead to enhanced skill level, thus, improvement in the metrics that are available for assessing leadership skills is warranted. Simulation-based training has been a useful tool for incorporating testing, training, and improvement in both technical and nontechnical skills. Future work will involve evaluating the relationship between directional instructions and hands-on procedural skills. To accomplish this, our laboratory is developing a validated grading system for final product score of the bowel repair, so we can related the nontechnical skill of directing to the technical skills required during the repair.

Acknowledgment

Funding for this study came from the Department of Defense grant #W81XWH-13-1-0080 entitled, "Psychomotor and Error Enabled Simulations: Modeling Vulnerable Skills in the Pre-Mastery Phase."

Authors' Contributions:

S.G. contributed to study conception and design, data collection, analysis and interpretation, writing the manuscript, and critical revision of the manuscript. K.L. contributed to study conception and design, data collection, analysis and interpretation, writing the manuscript, and critical revision of the manuscript. R.R. contributed to study data collection, analysis and interpretation, writing the manuscript, and critical revision of the manuscript. J.N. contributed to study data collection, analysis and interpretation, writing the manuscript, and critical revision of the manuscript. S.D. contributed to study data collection, analysis and interpretation, writing the manuscript, and critical revision of the manuscript. A.-L. D'A. contributed to study conception and

design, data collection, analysis and interpretation, writing the manuscript, and critical revision of the manuscript. C.P. contributed to study conception and design; data collection, analysis and interpretation; writing the manuscript, and critical revision of the manuscript.

Disclosure

The authors reported no proprietary or commercial interest in any product mentioned or concept discussed in this article.

REFERENCES

1. Suliman A, Klaber RE, Warren OJ. Exploiting opportunities for leadership development of surgeons within the operating theatre. *Int J Surg*. 2013;11:6–11.
2. Yule S, Parker SH, Wilkinson J, et al. Coaching non-technical skills improves surgical residents' performance in a simulated operating room. *J Surg Educ*. 2015;72:1124–1130.
3. Yule S, Flin R, Paterson-Brown S, Maran N. Non-technical skills for surgeons in the operating room: a review of the literature. *Surgery*. 2006;139:140–149.
4. Hull L, Arora S, Kassab E, Kneebone R, Sevdalis N. Observational teamwork assessment for surgery: content validation and tool refinement. *J Am Coll Surg*. 2011;212:234–243.
5. Reader TW, Flin R, Cuthbertson BH. Communication skills and error in the intensive care unit. *Curr Opin Crit Care*. 2007;13:732–736.
6. Blocker RC, Duff S, Wiegmann D, et al. Flow disruptions in trauma surgery: type, impact, and affect. *Hum Fac Erg Soc P*. 2012;56:811–815.
7. Wiegmann DA, ElBardissi AW, Dearani JA, Daly RC, Sundt TM. Disruptions in surgical flow and their relationship to surgical errors: an exploratory investigation. *Surgery*. 2007;142:658–665.
8. Catchpole K, Mishra A, Handa A, McCulloch P. Teamwork and error in the operating room - analysis of skills and roles. *Ann Surg*. 2008;247:699–706.
9. Pena G, Altree M, Field J, et al. Nontechnical skills training for the operating room: a prospective study using simulation and didactic workshop. *Surgery*. 2015;158:300–309.
10. Hull L, Arora S, Aggarwal R, Darzi A, Vincent C, Sevdalis N. The impact of nontechnical skills on technical performance in surgery: a systematic review. *J Am Coll Surg*. 2012;214:214–230.
11. Pugh CM, DaRosa DA, Santacaterina S, Clark RE. Faculty evaluation of simulation-based modules for assessment of intraoperative decision making. *Surgery*. 2011;149:534–542.
12. Scott DJ, Cendan JC, Pugh CM, Minter RM, Dunnington GL, Kozar RA. The changing face of surgical education: simulation as the new paradigm. *J Surg Res*. 2008;147:189–193.
13. Gordon JA, Wilkerson WM, Shaffer DW, Armstrong EG. "Practicing" medicine without risk: students' and educators' responses to high-fidelity patient simulation. *Acad Med*. 2001;76:469–472.
14. Pugh CM, Santacaterina S, DaRosa DA, Clark RE. Intra-operative decision making: more than meets the eye. *J Biomed Inform*. 2011;44:486–496.
15. DaRosa DA, Pugh CM. Error training: missing link in surgical education. *Surgery*. 2012;151:139–145.
16. D'Angelo ALD, Cohen ER, Kwan C, et al. Use of decision-based simulations to assess resident readiness for operative independence. *Am J Surg*. 2015;209:132–139.
17. D'Angelo ALD, Ray RD, Jenewein CG, Jones GF, Pugh CM. Residents' perception of skill decay during dedicated research time. *J Surg Res*. 2015;199:23–31.
18. Cooke JM, Larsen J, Hamstra SJ, Andreatta PB. Simulation enhances resident confidence in critical care and procedural skills. *Fam Med*. 2008;40:165–167.
19. Campbell J, Tirapelle L, Yates K, et al. The effectiveness of a cognitive task analysis informed curriculum to increase self-efficacy and improve performance for an open cricothyrotomy. *J Surg Educ*. 2011;68:403–407.
20. D'Angelo AL. *Evaluating operative performance through the lens of epistemic frame theory*. Masters [thesis]. Madison (WI): University of Wisconsin-Madison; 2015.
21. Flin R, Yule S, Rowley D, Maran N, Paterson-Brown S. *The non-technical skills for surgeons (NOTSS) system handbook v1.2: structuring observation, rating and feedback of surgeons' Behaviours in the operating theatre*. Aberdeen, Scotland: University of Aberdeen; 2006.
22. Flin R. The NOTECHS system. In: *Aviation psychology: Practice and research*. Aldershot, UK: Ashgate; 2004.
23. Parker SH, Yule S, Flin R, McKinley A. Surgeons' leadership in the operating room: an observational study. *Am J Surg*. 2012;204:347–354.
24. Sachs TE, Pawlik TM. See one, do one, and teach none: resident experience as a teaching assistant. *J Surg Res*. 2015;195:44–51.
25. Flin R, Yule S, Paterson-Brown S, Maran N, Rowley D, Youngson G. Teaching surgeons about non-technical skills. *Surgeon*. 2007;5:86–89.
26. Norgaard B, Ammentorp J, Kyvik KO, Kofoed PE. Communication skills training increases self-efficacy of health care professionals. *J Contin Educ Health Prof*. 2012;32:90–97.

The American Journal of Surgery (2016) 212, 573–578

The American Journal of Surgery

Association of Women Surgeons

Can a virtual reality assessment of fine motor skill predict successful central line insertion?

Hossein Mohamadipanah, Ph.D.[a], Chembian Parthiban, M.S.[b], Jay Nathwani, M.D.[a], Drew Rutherford, M.S.[c], Shannon DiMarco, B.A.[a], Carla Pugh, M.D., Ph.D.[a,]*

[a]*Department of Surgery, School of Medicine and Public Health, University of Wisconsin-Madison, 600 Highland Avenue, Madison, WI 53726, USA;* [b]*Department of Mechanical Engineering, College of Engineering and* [c]*Department of Kinesiology, School of Education, University of Wisconsin-Madison, Madison, WI, USA*

KEYWORDS:
Fine motor skills;
Assessment;
Subclavian central line;
Virtual reality

Abstract

BACKGROUND: Due to the increased use of peripherally inserted central catheter lines, central lines are not performed as frequently. The aim of this study is to evaluate whether a virtual reality (VR)–based assessment of fine motor skills can be used as a valid and objective assessment of central line skills.

METHODS: Surgical residents (N = 43) from 7 general surgery programs performed a subclavian central line in a simulated setting. Then, they participated in a force discrimination task in a VR environment. Hand movements from the subclavian central line simulation were tracked by electromagnetic sensors. Gross movements as monitored by the electromagnetic sensors were compared with the fine motor metrics calculated from the force discrimination tasks in the VR environment.

RESULTS: Long periods of inactivity (idle time) during needle insertion and lack of smooth movements, as detected by the electromagnetic sensors, showed a significant correlation with poor force discrimination in the VR environment. Also, long periods of needle insertion time correlated to the poor performance in force discrimination in the VR environment.

CONCLUSIONS: This study shows that force discrimination in a defined VR environment correlates to needle insertion time, idle time, and hand smoothness when performing subclavian central line placement. Fine motor force discrimination may serve as a valid and objective assessment of the skills required for successful needle insertion when placing central lines.

Evaluation of surgical skills is currently performed using human observation. For example, objective structured assessment of technical skills[1] is widely used in the surgical education community. Although objective structured assessment of technical skills and other methods are useful, these evaluations are labor-intensive[2] and require the use of surgical experts who have limited amounts of time for evaluation activities. In addition, observation-based assessments are subjective and prone to bias.[1] For instance, the halo effect can influence the evaluation of surgical skills.[3]

Funding for this study came from the Department of Defense grant #W81XWH-13-1-0080 entitled “Psycho-Motor and Error Enabled Simulations Modeling Vulnerable Skills in the Pre-Mastery Phase–Medical Practice Initiative Procedural Skill Decay and Maintenance (MPI-PSD)”.

The authors declare no conflicts of interest.

* Corresponding author. Tel.: +1-608-262-2147; fax: +1-608-263-2354.

E-mail address: pugh@surgery.wisc.edu

Manuscript received February 1, 2016; revised manuscript June 29, 2016

0002-9610/$ - see front matter © 2016 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amjsurg.2016.06.023

In an attempt to address the limitations of observer-based feedback, there is an increase in the amount of research dedicated to nonobserver-based (objective) evaluations of surgical skills.[4–6] Advances in sensor technologies and the computer sciences enable this effort and provide a potential means of achieving automated, objective evaluations. For example, a variety of motion metrics have been developed to assess surgical performance by tracking the motion of surgical tools.[7] Also, virtual reality (VR) has been used to educate and assess trainees' psychomotor skills outside the operating room.[8]

Assessment of psychomotor skills has been conducted by defining motion metrics that have correlations to surgical skills.[7] For instance, in the assessment of predefined tasks on a laparoscopic VR simulator, variables such as smoothness, path length, and task execution time have been validated as objective assessment metrics.[9] In another study, a laparoscopic VR simulator was validated as an objective assessment tool assessing fine motor metrics using angular path length[10] as a measure of skill. Although previous studies have demonstrated validity evidence for objective measures of surgeons' psychomotor skills, a finer relationship between objective metrics and specified procedure steps would provide more explicit guidelines for training and remediation. Most of the research on motion metrics has focused on motor skills and not decision-making, the latter of which is an important part of surgical performance.[11] This is the first time that the metric idle time is used to investigate decision-making skills in the subclavian central venous catheter (CVC).

The aim of this study is to identify a relationship between the ability of surgical residents to discriminate forces in a VR environment and the use of specific motor movements during needle insertion for subclavian CVC. We hypothesize that ability in fine motor discrimination, as assessed in a VR environment will have significant correlations with the gross motor movements during subclavian CVC.

Methods

Participants

Residents from 7 general surgery programs in the United States (N = 43) participated in this study. All participants identified themselves as right handed, with exception to 1 participant who was left handed and 1 who was ambidextrous. The University of Wisconsin Hospitals and Clinics Institutional Review Board has approved this study.

Motion tracking system and motion metrics

Each participant was equipped with a head camera that recorded his or her performance during the placement of the subclavian CVC. They were additionally equipped with motion-tracking sensors from a Motion Monitor system (Innovative Sport Training, Inc), as shown in Fig. 1. These magnetic sensors were affixed to participants' index fingers, thumbs, and wrists in standard positions. The motion tracking system uses sensors that provide 6 variables, which include the position and orientation of each sensor. Fig. 2 shows examples of position and orientation relative to hand movements. The motion data received by the sensors is filtered by a low-pass second-order Butterworth filter (with a cutoff frequency of 1 Hz[12]) to remove noise.

We used a total of 4 metric points of interest from the hand movement data while participants were performing the subclavian CVC insertion. "Insertion time" is defined as the duration of needle insertion into the manikin. "Idle time"[4] is defined as the duration of time that the magnitude of velocity of all sensors was less than .02 m/s when executing needle insertion. "Position smoothness" is based on the derivative of the position acceleration,[13] and similarly "orientation smoothness" is based on the derivative of the orientation acceleration. Greater magnitudes for the smoothness metrics represent greater smoothness.

Subclavian central line simulation

Participants were asked to insert a subclavian central line on a simulated hypotensive, tachycardic, and febrile patient. A complete subclavian CVC kit with all the necessary equipment was provided to the participants to perform a subclavian CVC insertion. The simulation model (CentralLineMan System, Simulab Corp, Seattle, WA) had a realistic anatomy of a right upper adult torso. The simulator was draped and prepped for the procedure in a standard surgical fashion. The simulator was placed in Trendelenburg position (24°). The participant's view of the scenario is shown in Fig. 3 through a head camera. Our particular interest was in needle insertion duration, as proper needle insertion is a crucial subtask for successful subclavian central line placement.[14] Hence, all the hand

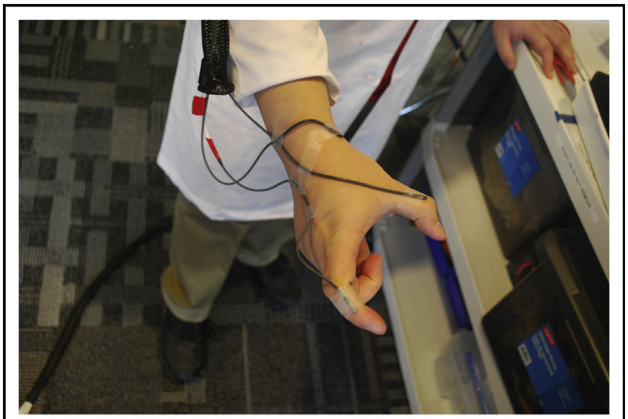


Figure 1 Sensors attached to a participant's index finger, thumb, and wrist.

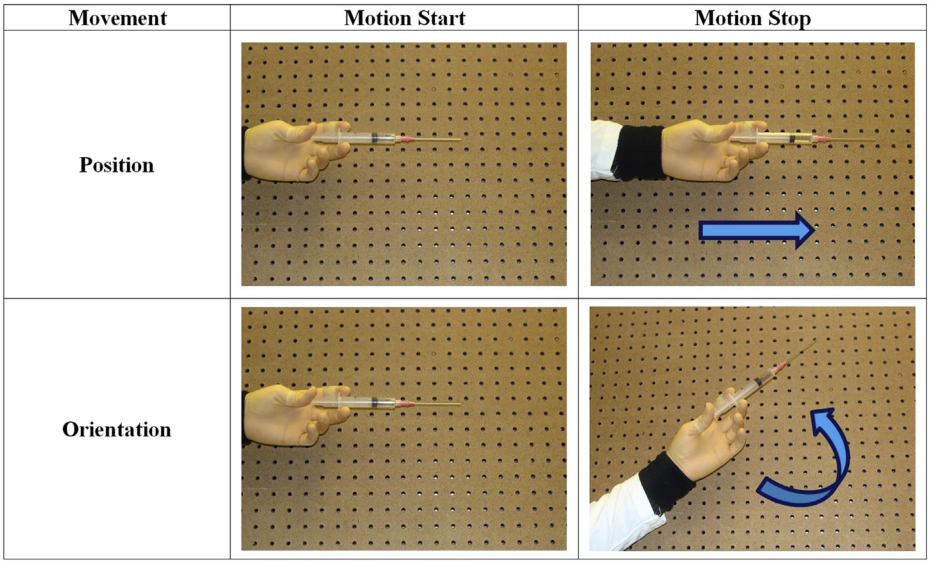


Figure 2 Two types of hand movement; (top) position movement (bottom) orientation movement.

movement metrics for this study are calculated only during needle insertion.

Fine motor force discrimination using virtual reality

As an additional task, the participants were asked to perform a fine motor force discrimination test with the help of a VR haptic device (Force Dimension Omega) as shown in Fig. 4. The haptic device provides force feedback based on interactions of the operator within a VR environment. The device includes a virtual stylus that can be moved by a participant in the 3-dimensional environment to feel the reaction force of 2 predefined blocks, namely blocks A and B in Fig. 5. Participants are asked to discriminate forces within the 2 blocks by tapping on them and feeling the force feedback given by the haptic device.

Each trial had a time limit of either 30, 20, or 15 seconds to force participants to make a decision regarding their comparison of force for the 2 blocks. The initial 2 trials were set to a 30-second time limit to make sure the participant understands the task. Subsequent trials were set to have 15, 20 or 30 seconds time limits. We have provided an Appendix (Table 1) with the exact sequencing. Once the participant sampled the force discrimination, he/she indicated which block had the highest force by clicking the button on the stylus while the stylus was placed on the respective block. Some trials had blocks with equal forces. In this case, the participant clicked on a designated area between the blocks. Successful discrimination (correctness) was displayed on the screen after every trial and was recorded for offline analysis. Additional details for the force discrimination task in the VR environment are provided in the Appendix. In all trials, the total number of correct force discriminations was labeled “correct discrimination”, whereas the total number of wrong force discriminations was labeled “wrong discrimination”. To make sure that each participant understood the task, the first 5 trials of each participant were excluded from analysis.

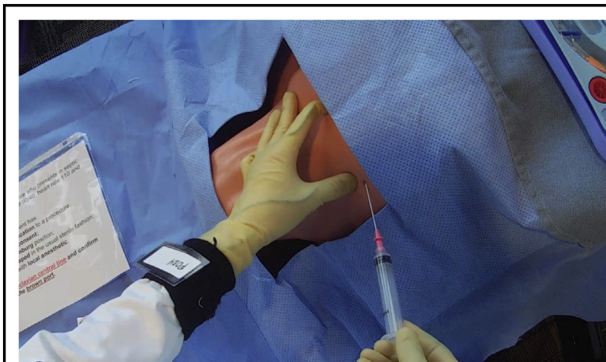


Figure 3 The participant’s view of the subclavian central venous catheter scenario.

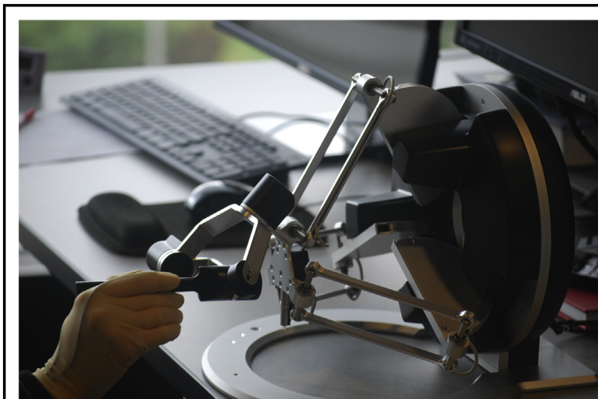


Figure 4 The haptic device operated by a participant during data collection.

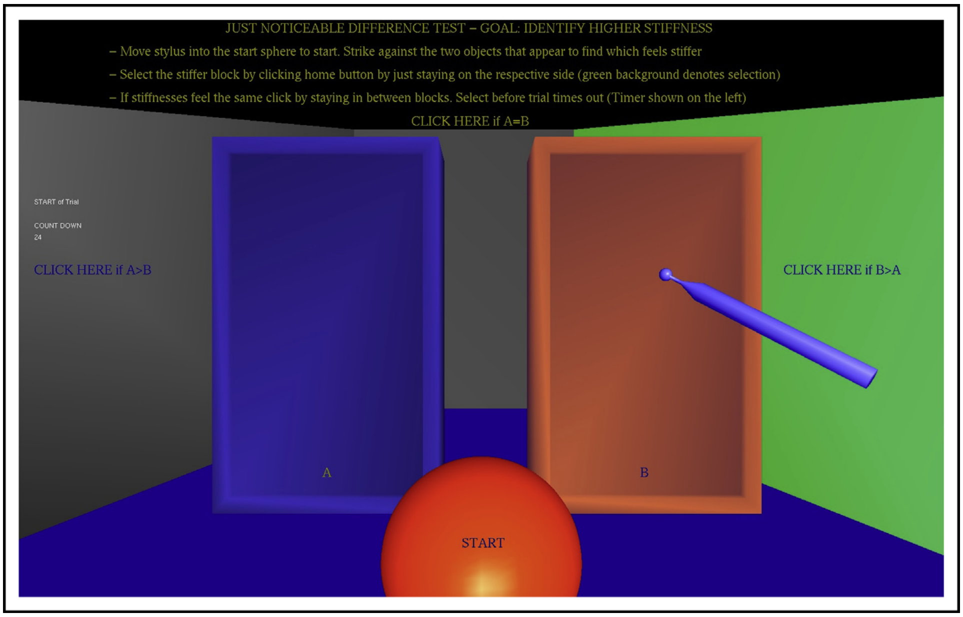


Figure 5 Screenshot of the force discrimination task in the virtual environment. Shows blocks A and B of different forces being rendered and the virtual stylus (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Results

Table 2 illustrates the correlation between VR force discrimination metrics and the time-related metrics in subclavian CVC. The metric "correct discrimination" has a negative correlation ($r = -.3802$, $P = .0119$), and the metric "wrong discrimination" has a positive correlation ($r = +.3317$, $P = .0298$) to the "insertion time" metric meaning, persons with better fine motor skills had less insertion time. In addition, the metric "correct discrimination" has a negative correlation ($r = -.4628$, $P = .0018$) to the "idle time" metric, and the metric "wrong discrimination" has a positive correlation ($r = +.4108$, $P = .0062$) to the "idle time" metric. In essence, those persons with excellent fine motor force discrimination have less pauses during needle insertions, and those with poor fine motor force discrimination have longer pauses. All the reported correlation coefficients in this study are Spearman's rho coefficients.

Videos of subclavian CVC indicated that all participants used their right hand for handling the syringe including the left handed and ambidextrous participants. This syringe handling may be the result of simulation scenario setting where the head of the simulator is located on the left side of the participant, shown in Fig. 3. There was not an option to stand on the other side. To assess for potential effects, we calculated the average of the fine motor metrics from the 3 sensors on the left hand and the average of the fine motor metrics from the 3 sensors on the right separately.

Table 3 presents the calculated correlations between VR force discrimination metrics and the position and orientation smoothness in subclavian CVC. The metric "correct discrimination" in the VR environment demonstrates positive correlation to the metrics "position smoothness left hand" ($r = +.3797$, $P = .012$), "position smoothness right hand" ($r = +.3683$, $P = .0151$), "orientation smoothness left hand" ($r = +.3269$, $P = .0324$), and "orientation smoothness right hand" ($r = +.4135$, $P = .0059$) from subclavian CVC simulation. In addition, the metric "wrong discrimination" in the VR environment did not correlate ($P > .05$) to any metric from subclavian CVC simulation.

We set $P < .05$ as the significant level in our analysis; however, to have a more conservative analysis, we used Bonferroni importance level $P<\left(\frac{0.05}{6}=0.0083\right)$ (as there are 6 motion metrics in this study). By considering this significant level in Table 1, the metric "correct discrimination" has a negative significant correlation, and the metric "wrong discrimination" has a positive significant correlation ($P < .0083$) to the "idle time" metric. In Table 2, the metric "correct discrimination" has shown significant positive correlation ($P < .0083$) to the metrics "orientation smoothness right hand".

Table 1 The correlation between force discrimination metrics in VR and the time-related metrics in subclavian CVC

Metric of subclavian central line	Metric of force discrimination in VR	R	P value
Insertion time	Correct discrimination	−.3802	.0119
	Wrong discrimination	+.3317	.0298
Idle time	Correct discrimination	−.4628	.0018
	Wrong discrimination	+.4108	.0062

CVC = central venous catheter; VR = virtual reality.

Comments

This study sought to validate the relationship between fine motor force discrimination in a VR environment and gross

Table 2 The correlation between the fine motor metrics in subclavian CVC and VR force discrimination metrics

Metric of subclavian central line	Metric of force discrimination in VR	*R*	*P* value
Position smoothness Left hand	Correct discrimination	+.3797	.0120
	Wrong discrimination	−.2568	.0965
Position smoothness Right hand	Correct discrimination	+.3683	.0151
	Wrong discrimination	−.2113	.1738
Orientation smoothness Left hand	Correct discrimination	+.3269	.0324
	Wrong discrimination	−.2024	.1931
Orientation smoothness Right hand	Correct discrimination	+.4135	.0059
	Wrong discrimination	−.2238	.1492

CVC = central venous catheter; VR = virtual reality.

motor metrics in the insertion of a subclavian CVC on a simulator. Our analysis reveals that participants who had less insertion time made a higher number of correct decisions and lower number of wrong decisions during the VR discrimination task. Also, correlations between the metric "idle time" and the "correct discrimination" and "wrong discrimination" in Table 1 implies that participants with more ability in discriminating forces in the VR environment had less amount of uncertainty during their needle insertion. The findings in Table 2 show that smoother position and orientation movements of the hands imply a certain level of underlying proficiency in needle insertion ability during subclavian CVC. As all the participants held the syringe with their right hands, the findings in Tables 1 and 2 imply that orientation smoothness of the syringe-holding hand and "idle time" are the most informative ($P < .0083$) fine motor metrics in subclavian CVC procedure.

Findings in this study also indicate that haptic sensation is an important motor skill that could affect performance of needle insertion in subclavian CVC. Those who have poor haptic sensation may miss important cues when passing through the chest wall tissues. For instance, when the needle first hits the clavicle, the participant should put light pressure on the needle with their opposite hand for guidance of the needle into the subclavian vein. Those who fail to feel the needle hit the clavicle may continue to exert force, causing tissue damage and undue pain for the patient. In addition, these actions may result in an unsuccessful venous puncture.

Moreover, we are going to study the effect of haptic sensation skills on the other open and laparoscopic surgical scenarios, including anastomosis bowel repair and laparoscopic ventral hernia repair, in our future studies. In addition, VR simulators, such as the minimally invasive surgical trainer, virtual reality,[15] have been widely used and studied in surgical education and have been successful in showing that training in a virtual environment can reduce errors in the operating room. Consequently, there is an opportunity to use VR to test specific surgical and psychomotor skills, as these systems can mimic a wide range of haptic sensations and automatically record detailed performance metrics based on human computer interaction.

Although "smoothness" and "idle time" metrics have shown great potential as motion metrics in evaluating subclavian CVC, more work needs to be done to translate these metrics into learning objectives and skill exercises for trainees. For example, telling a resident that their performance was not smooth enough is not specific enough to improve performance.

There are several other limitations in this study. One is the small sample size that we plan to address by conducting additional data collections. The other limitation was the

Table 3 The summary of the trial condition parameters and its frequency distribution as used for the 43 participants

Trial condition	Time limit (s)	Force of block A (N/mm)	Force of block B (N/mm)	Force difference (N/mm)	Combo 1 frequency	Combo 2 frequency	Combo 3 frequency
1	30	8.700	11.600	2.900	2	2	2
2	20	10.150	5.800	4.350	1	1	0
3	30	7.250	4.350	2.900	1	1	1
4	30	5.800	8.700	2.900	1	2	1
5	15	8.700	8.700	.000	1	1	2
6	15	4.350	2.900	1.450	1	2	3
7	20	13.050	13.050	.000	1	1	0
8	15	2.175	2.175	.000	1	1	0
9	20	4.350	11.600	7.250	1	1	4
10	15	13.050	5.800	7.250	2	1	1
11	20	2.900	1.450	1.450	2	1	2
12	15	1.450	2.900	1.450	1	1	1
13	20	2.175	3.625	1.450	1	1	1
14	15	6.525	7.250	.725	1	1	0
15	20	1.450	10.150	8.700	1	1	0
16	15	11.600	12.325	.725	1	1	1

The 19 trials were a combination of the 16 trial conditions that were randomly assigned to the participants in the form of 3 combination files, called combo 1, combo 2, and combo 3. Each of combo 1 and combo 2 was used for 15, and combo 3 was used for 13 participants of total 43 participants.

nature of participation, which was on a voluntary basis of the general surgery residents. Therefore, the results of this study exclude the residents who did not accept to be a volunteer and participate in this study. Future work will also focus on assessing the difference between the VR force discrimination metrics and global ratings obtained by experts for subclavian CVC. This will allow additional validation of our findings and an assessment of how VR assessments could be used to guide fine motor skills training for residents as faculty time is limited and may be less objective and instructive.

Supplementary Data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.amjsurg.2016.06.023.

References

1. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg 1997;84:273–8.
2. Sidhu RS, Grober ED, Musselman LJ, et al. Assessing competency in surgery: where to begin? Surgery 2004;135:6–20.
3. Wanzel KR, Ward M, Reznick RK. Teaching the surgical craft: from selection to certification. Curr Probl Surg 2002;39:583–659.
4. D'Angelo A-LD, Rutherford DN, Ray RD, et al. Idle time: an underdeveloped performance metric for assessing surgical skill. Am J Surg 2015;209:645–51.
5. Dosis A, Aggarwal R, Bello F, et al. Synchronized video and motion analysis for the assessment of procedures in the operating theater. Arch Surg 2005;140:293–9.
6. Datta V, Mandalia M, Mackay S, et al. Relationship between skill and outcome in the laboratory-based model. Surgery 2002;131:318–23.
7. Oropesa I, Sánchez-González P, Lamata P, et al. Methods and tools for objective assessment of psychomotor skills in laparoscopic surgery. J Surg Res 2011;171:e81–95.
8. Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. Br J Surg 2004;91:1549–58.
9. Pellen MGC, Horgan LF, Barton JR, et al. Construct validity of the ProMIS laparoscopic simulator. Surg Endosc 2009;23:130–9.
10. Dongen KW, Tournoij E, Zee DC, et al. Construct validity of the LapSim: can the LapSim virtual reality simulator distinguish between novices and experts? Surg Endosc 2007;21:1413–7.
11. Pugh C, Plachta S, Auyang E, et al. Outcome measures for surgical simulators: is the focus on technical skills the best approach? Surgery 2010;147:646–54.
12. Falk V. Manual control and tracking—a human factor analysis relevant for beating heart surgery. Ann Thorac Surg 2002;74:624–8.
13. Hogan N, Sternad D. Sensitivity of smoothness measures to movement duration, amplitude and arrests. J Mot Behav 2009;41:529–34.
14. McGee DC, Gould MK. Preventing complications of central venous catheterization. N Engl J Med 2003;348:1123–33.
15. Maithel S, Sierra R, Korndorffer J, et al. Construct and face validity of MIST-VR, Endotower, and CELTS. Surg Endosc 2006;20:104–12.

Appendix

In the virtual reality environment, each participant was provided with 19 trials, where the object forces were set between (1.45–13.05) N/mm which corresponds to 10% and 90% of the maximum closed loop force handled by the device, which is 14.5 N/mm. The object force selections were based on a randomly repeated combination of 16 different trial conditions, which is shown in Table 3. The effective force discrimination between objects A and B shown in Fig. 5 was between (0 and 8.7) N/mm, where 0 means the objects had equal force, and 8.7 N/mm was the maximum difference in force among all the trials.

Kimberly Maciolek (1 month)
Martin Moe (6 months)
Bridget O'Connell-Long (6 months)
Katherine Peterson (3 months)
Philip Terrien (6 months)
Aliyya Terry (3 months)
Nicole Van Beek (9 months)
Anna Witt (18 months)

Relationship Between Technical Errors and Decision-Making Skills in the Junior Resident

Jay N. Nathwani, MD,* Rebekah M. Fiers,* Rebecca D. Ray, PhD,* Anna K. Witt,* Katherine E. Law, MS,† ShannonM. DiMarco, BA,* and Carla M. Pugh, MD, PhD*

*Division of General Surgery, University of Wisconsin, Madison, Wisconsin; and †Department of Industrial and Systems Engineering, University of Wisconsin, Madison, Wisconsin

OBJECTIVE: The purpose of this study is to coevaluate resident technical errors and decision-making capabilities during placement of a subclavian central venous catheter (CVC). We hypothesize that there would be significant correlations between scenario-based decision-making skills and technical proficiency in central line insertion. We also predict residents would face problems in anticipating common difficulties and generating solutions associated with line placement.

DESIGN: Participants were asked to insert a subclavian central line on a simulator. After completion, residents were presented with a real-life patient photograph depicting CVC placement and asked to anticipate difficulties and generate solutions. Error rates were analyzed using chi-square tests and a 5% expected error rate. Correlations were sought by comparing technical errors and scenario-based decision-making skills.

SETTING: This study was performed at 7 tertiary care centers.

PARTICIPANTS: Study participants ($N = 46$) largely consisted of first-year research residents who could be followed longitudinally. Second-year research and clinical residents were not excluded.

RESULTS: In total, 6 checklist errors were committed more often than anticipated. Residents committed an average of 1.9 errors, significantly more than the 1 error, at most, per person expected ($t(44) = 3.82$, $p < 0.001$). The most common error was performance of the procedure steps in the wrong order (28.5%, $p < 0.001$). Some of the residents (24%) had no errors, 30% committed 1 error, and 46 % committed more than 1 error. The number of technical errors committed negatively correlated with the total number of commonly identified difficulties and generated solutions ($r\ (33) = -0.429$, $p = 0.021$, $r\ (33) = -0.383$, $p = 0.044$, respectively).

CONCLUSIONS: Almost half of the surgical residents committed multiple errors while performing subclavian CVC placement. The correlation between technical errors and decision-making skills suggests a critical need to train residents in both technique and error management. (J Surg Ed 73:e84-e90. © 2016 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: surgical education, simulation, performance assessment

COMPETENCIES: Medical Knowledge, Practice-Based Learning and Improvement, Systems-Based Practice

INTRODUCTION

Central venous catheter (CVC) placement is a common procedure in the intensive care unit setting. In the United States, more than 5 million catheters are placed on a yearly basis.[1] Common indications for central line placement include administration of vasoactive medications, rapid resuscitation, total parenteral nutrition, and delivery of caustic medications.[2] When considering the options for central venous access, placement of a subclavian CVC is often preferred because of its low infection rate.[3]

Though considered a simple procedure, central line insertion is not without risk. It is estimated that up to 15% of patients who undergo central line insertion would be confronted with at least 1 complication, including infection, arterial puncture, pneumothorax, hemothorax, mediastinal hematoma, and vascular thrombosis.[1,3-5] Data show that experience and proper training are crucial to minimize such complications.[5,6] At tertiary care centers, residents are often responsible for inserting CVCs.[6] Because

Correspondence: Inquiries to Carla M. Pugh, MD, PhD, FACS, Department of Surgery, University of Wisconsin Hospital and Clinics, 600 Highland Avenue, WiSOR K6/135, Madison, WI 53792; e-mail: pugh@surgery.wisc.edu

  1931-7204/$30.00
http://dx.doi.org/10.1016/j.jsurg.2016.08.004

resident experience can be highly variable, education is incredibly important to help minimize complications.[6,7] In addition, confirmation of adequate resident skill is in critical need.[8]

Resident education continues to improve with goals of decreasing patient morbidity. In central line training, simulation helps house staff members develop baseline procedural skills before working on patients.[8,9] Other efforts at decreasing resident errors include video review of real-time performance.[10] Although valuable in providing important technical assessment, these studies are limited by providing an incomplete assessment of CVC insertion. Recent studies have improved the process of assessment, recognizing the importance of decision making and encouraging the incorporation of broader assessment methods.[11]

The purpose of this study is to improve on current resident assessment methods and evaluate both resident technical errors and decision-making capabilities during placement of a subclavian CVC in a simulated environment. We predict that there would be significant correlations between scenario-based decision-making skills and technical proficiency in central line insertion. In addition, we predict that residents would face some problems in anticipating common difficulties and generating solutions associated with line placement.

MATERIALS AND METHODS

Setting and Participants

As a part of a longitudinal skills decay study, general surgery programs in the Midwestern United States were sent inquiries for participation. Recruitment efforts were undertaken by phone or by electronic correspondence to either program directors or resident education coordinators. Overall, 7 programs agreed to participate. Primary participant recruitment efforts focused on first-year research residents; however, clinical residents and research residents in their second year who expressed interest in participation were not excluded. Data collection occurred at the respective training program's location during the summer of 2014, and all participation was voluntary. This study was approved by the University of Wisconsin Hospital and Clinics (UWHC) Institutional Review Board.

Research Protocols

Participants first completed a demographic survey. Once completed, they were guided to the central line station. A research assistant read a standardized introduction to the central line task.

Subclavian Central Line Insertion

Participants were provided a clinical scenario of an elderly male patient in septic shock in need of a central line. Participants were asked to insert a subclavian CVC on a central line simulator that has a realistic anatomical representation of the right upper torso of an adult male, including the subclavian, internal, and external jugular veins, and carotid and subclavian arteries. Arterial blood was represented with red liquid and venous blood by blue liquid. Although there is an evolving standard for ultrasound-guided internal jugular central lines, subclavian was chosen because the simulator used was not compatible with ultrasound. As such, both subclavian and internal jugular accesses were viewed as equally appropriate options.[3]

The patient was presumed to have given consent, have no contraindication to the procedure, be placed in Trendelenburg position, be prepared and draped in the usual fashion, and be already injected with local anesthetic. Participants were encouraged to verbalize any steps they would perform normally but could not complete because of the limitation of the simulated environment.

Participants were provided with a full central line insertion kit and 15 minutes to complete the procedure portion of the task. Owing to limitation of the simulator, a scalpel handle without a blade was provided to participants who were asked to still go through the motion of scalpel usage to represent the order in which it would have been used.

Subclavian Cognitive Scenarios

When the simulation was completed, participants were directed to the second portion of the station where they were presented with a cognitive scenario. The scenario was a morbidly obese 36-year-old female patient, postoperative day 2, in septic shock, and in need of transfer to the intensive care unit for vasoactive therapy. Participants were provided with a picture of the patient's upper torso (Fig. 1). Participants were then asked to describe any potential difficulties they anticipated given the scenario and provide solutions to the difficulties when placing a subclavian central line.

Data Analysis

Study data were collected and managed using REDCap electronic data capture tools.[12] REDCap is a secure, web-based application designed to support data capture for research studies. Data collection focused on a checklist of technical errors of varying morbidity (Appendix). The checklist used to track participant technical errors was adopted based on a literary review of the most commonly committed errors in subclavian central line insertion. The checklist was reviewed by an expert for final approval. Verbal responses to the cognitive scenarios were transcribed. Anticipated difficulties and solutions were grouped by similarity in theme. As an example, if participants listed multiple positioning maneuvers to optimize subclavian

FIGURE 1. Patient picture presented to participants for cognitive scenario. Participants were asked to anticipate difficulties and generate solutions when considering insertion of a subclavian CVC. (https://psnet.ahrq.gov/webmm/case/221, Accessed 16.03.16).

CVC placement as a solution, the responses were grouped as positioning maneuvers.

Data analyses were performed in SPSS 23.[13] Descriptive statistics and chi-square analyses were performed to evaluate technical error rates. Using a 5% error rate as the expected, chi-square analyses were performed on all errors. Pearson correlations were performed between the total number of errors on the central line and the number of commonly identified difficulties and solutions given in the central line scenario.

RESULTS

A total of 46 general surgery residents (53.2% female) performed this task. Participants were between their second and fourth postgraduate year. Furthermore, 47% of residents were in their first research year, 36.2% were in their second research year, and 14.9% were in their second through fourth clinical years. The resident's years of clinical experience varied: 3 participants completed 1 clinical year, 32 participants completed 2 clinical years, 10 participants completed their third clinical year, and 1 completed his/her fourth clinical year.

Subclavian Central Line Insertion

Participants demonstrated poor needle angle more often than expected (26.7% of participants, $\chi^2_1 = 44.47$, $p < 0.001$). Participants also redirected the needle without adequate withdrawal more than expected (15.6% of participants $\chi^2_1 = 42.19$, $p < 0.001$). Less than 2 attempts to access the vein with the needle was coded as acceptable; however, more than 22% of participants took more than 2 attempts to access the vein, with a maximum of 7 attempts ($M = 2.04$, standard deviation $= 1.77$, $\chi^2_1 = 28.10$, $p < 0.001$). Overall, 13.6% of participants required more than 1 guidewire ($\chi^2_1 = 6.91$, $p < 0.009$). A greater than expected number of participants forgot to use the scalpel

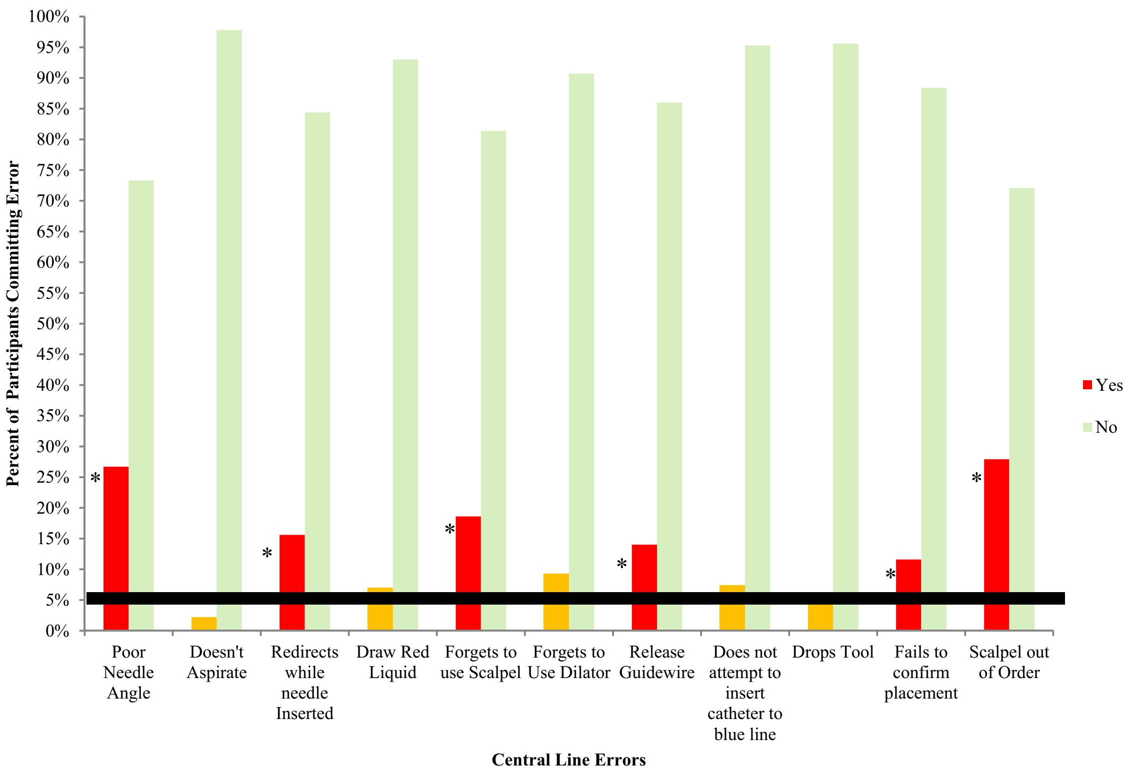


FIGURE 2. Percent of participants performing central line errors. Solid black line represents expected error rate, used for chi-square analysis. *Errors committed in a statistically significant fashion ($p < 0.05$). Yellow bars represent errors that were committed, but did not achieve statistical significance.

(18.6%, $\chi_1^2 = 16.76$, $p < 0.001$), and 13.9% of participants dropped the guidewire on the ground during the procedure ($\chi_1^2 = 7.26$, $p < 0.007$). Furthermore, 11.6% failed to confirm placement of the catheter, and 28.5% of the participants performed the steps of the central line procedure in the wrong order ($\chi_1^2 = 3.98$, $p < 0.046$ and $\chi_1^2 = 47.82$, $p < 0.001$, respectively) (Fig. 2).

Current assessment literature does not describe standard error expectation rates for technical performance of subclavian CVC placement. The American College of Surgeons/Association for Program Directors in Surgery Residents Skills (ACS/APDS) curriculum lists central line insertion a phase 1 skill.[14] Based on the expectation set by this curriculum that a surgical resident should competently place a central line while in junior training, we used 1 error per participant as our expected error rate. Participants averaged a greater than expected 1.87 errors (standard deviation = 1.5, t (44) = 3.82, $p < 0.001$). A minimum number of participants committed no errors (24%); 30% committed 1 error. Moreover, 46 % committed more than 1 error, and 2% committed a maximum of 6 errors (Fig. 3).

Subclavian Cognitive Scenarios

The 3 most common anticipated difficulties reported by participants for the cognitive scenario included difficulty associated with the patient's obese body habitus (63%), difficulty in identifying land marks (48%), and dense soft tissue (35%). The 3 most common proposed solutions to overcome the anticipated difficulties included use of ultrasound (41%), placing an internal jugular central line instead of a subclavian (41%), and positioning maneuvers (24%).

The Relationship Between Technical Errors and Cognitive Performance

The number of total errors committed during the subclavian central line procedure negatively correlated with the number of possible difficulties that residents could list while placing a central line (r (33) = −0.419, $p = 0.021$). The number of total errors showed a significant negative correlation with the number of top solutions that residents could generate in response to anticipated difficulties in placing a subclavian central line (r (33) = −0.383, $p = 0.044$). There was no correlation between number of completed clinical years, technical errors, and cognitive performance ($p > 0.05$).

DISCUSSION

This study examines surgical resident's technical skill and decision-making capacity while placing a subclavian venous catheter. We predicted there would be significant correlations between scenario-based decision-making skills and technical proficiency in central line insertion. In addition, we predicted that residents would face some problems in anticipating common difficulties and generating solutions associated with line placement. Our analyses indicate that most residents commit at least 1 cognitive or technical error while placing a subclavian venous catheter. The most common error committed by residents was performing the procedure in the wrong step order, performed by 28.5% of the participants. Additionally, there were 2 significant negative correlations between the number of errors committed in line insertion and (1) the number of anticipated difficulties, and (2) proposed solutions in difficult line placement. In other words, participants who were

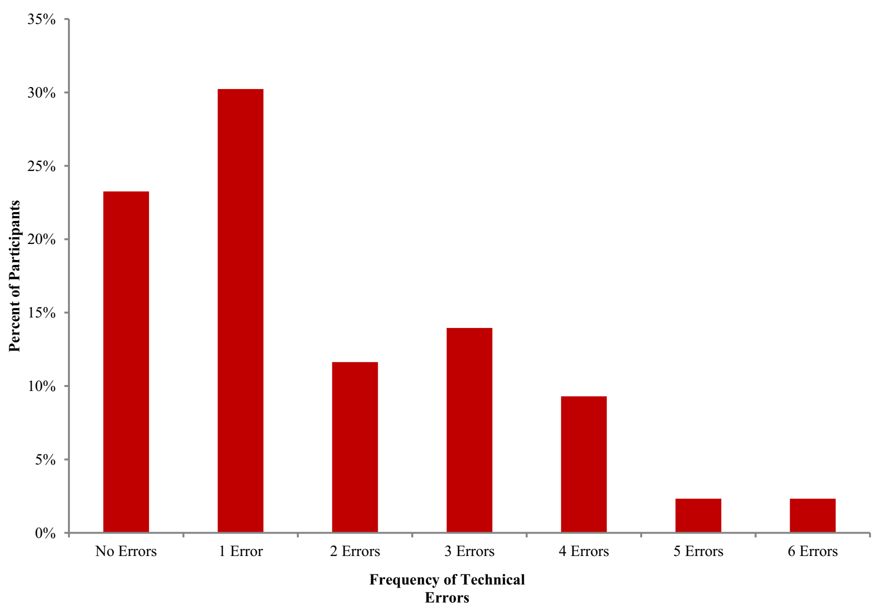


FIGURE 3. Percentage of participants committing frequency of technical errors.

more prepared to discuss difficulties and solutions in the cognitive scenario also committed a lower number of procedural errors during CVC placement. We believe that our data are a reflection of our participant's clinical experience in central line insertion.

Junior Resident Technical Performance in Subclavian CVC

The cognitive scenario was meant to illustrate a difficult central line scenario. As the presenting patient was morbidly obese, we anticipated participants would list the patient's obese habitus, increased soft tissue, and difficulty in identifying land marks as the top issues.[15] The increased volume of soft tissue would obscure landmarks and increase the difficulty of inserting the central line needle in the proper location and direction.[15] Additionally, the larger body habitus would likely translate to a deeper needle insertion, which would leave participants familiar with CVC insertions on an average-sized individual unable to rely on muscle memory while initially inserting the needle. Participants' top responses were in line with our expectations.

Most participants committed at least 1 error. With central line insertion listed as a phase 1 basic/core skill by the resident skills curriculum,[14] we presumed that most residents would have committed no errors. This unexpected finding could be a disconnect between curricular expectation and current practices. With the introduction of peripherally inserted central catheter (PICC) insertion and Interventional Radiology teams, the option to consult other services for central venous access may be a more common option during real-life scenarios. This could translate into fewer opportunities for junior residents to practice central line insertion in early training, leading to a lack of mastery and committing more procedural errors.

A potential solution for the training gaps and limited exposure to central line insertion could be through simulation. Clinical practice in simulation has led to increased procedural performance and improved patient safety.[8,16,17] Practicing in simulation could allow for multiple opportunities to train in a protected environment multiple times without harming a patient. Moreover, simulation could also be used after competency is achieved to test difficult scenarios and maintain baseline skills. The ACS/APDS surgical skills curriculum already provides the blueprint curriculum and could be further built on with cognitive scenarios and error checklists to offer a more comprehensive assessment.[14] Training centers could further use this developed curriculum to identify gaps in training, ensure competency, and maintain skills for procedures that may not be as often performed.

Junior Resident Cognitive Performance in Subclavian CVC

There was a significant negative correlation between the number of errors performed on the subclavian central line insertion and the anticipated difficulties and proposed solutions with our cognitive scenarios. Participants with more errors on the procedural portion of the study have less anticipated difficulties and solutions with the cognitive scenarios. This is likely related to experience with central line insertions. Several studies show that more experienced residents and clinicians have higher success and fewer complications when inserting CVCs relative to those less experienced.[6,7] This may be because central line insertion opportunities allow for greater exposure to challenges, allowing for mentored guidance and troubleshooting. These exposures and guidance likely translates into improved performance and ability to troubleshoot difficult catheter scenarios independently.

Implications for Surgical Education

Transformation of surgical education and concerns for patient safety necessitates novel teaching and assessment methods in training programs. This would facilitate competency in a limited period of time for both bedside and operating room procedures.[8] Examples of such methods in central line insertion include simulation practice and post-performance video review of technical aspects.[8,10] Recent studies indicate that assessment of cognitive skills in conjunction with technical skills is important for comprehensive assessment of procedural performance.[11] Our study is one of the first published literature that attempts to simultaneously understand cognitive and technical performance of the subclavian central line insertion in the resident population. Exploring performance by this method appreciates how these skills vary and relate by differences in experience level.

Our findings should be further expanded on in future works for improvement of surgical education. With our primary purpose being to coevaluate technical and decision-making skills, we wanted to show that simple methods could be developed and used to successfully achieve a more comprehensive assessment of surgical performance. We suggest that current, classic technical rating scales include evaluation of decision-making skills and other complex errors. With continued evolution in clinical assessment, it is possible to develop an objective method of technical and cognitive assessment—a goal that is strongly sought after in the field of surgery[11,18]

While achieving the objectives of the study, some limitations were apparent. Our scoring checklist assesses technical aspects of the subclavian central line procedure; however, it does not appreciate the magnitude that an error could contribute to patient morbidity. Exploring the weight of committed errors would allow us to better appreciate the relationship between high magnitude errors and postgraduate year, if applicable. The lack of statistical correlation between clinical experience with technical or cognitive skills is surprising. Intuitively, the expectation would be that clinical years of experience would directly relate to improved performance of central line insertion. Yet, we believe that a

possible reason for the lack of statistical significance could be the sample size and breadth of sample. Although we primarily focused on recruiting junior level residents, future work should evaluate resident performance with a larger, more diverse population.

CONCLUSIONS

This study suggests that junior residents currently have lower than expected performance in subclavian CVC insertion. Our results indicate current trainees are not gaining the necessary experience needed to achieve competency. There is a need for increased opportunities that develop subclavian CVC insertion mastery. Improvement in surgical education programs by use of simulation-based assessments that focus on technical and decision-making skills may help to increase successful performance.

ACKNOWLEDGMENTS

A total of 2 grants funded this study. The first is from the U.S. Department of Defense, United States (W81XWH-13-1-0080) grant. The second is from a National Institutes of Health, United States (T32CA090217) grant. Neither funding source contributed to the study design, data acquisition, analysis or interpretation, or decision to submit this article for publication.

APPENDIX A

STATION A COMMON EVENTS CHECKLIST

Directions: For each participant mark [X] any of the items that occur during the procedure (i.e. mark [X] in rows 1 and 2 below "Poor needle angle" if the participant does this twice).

Participant ID: ______________ Researcher Initials: __________ Date: ______________

	Event #	Common Events That Can Occur													
		Poor needle angle	Fail to aspirate with needle insertion	>3 attempts to access the vessel	Draw "red" liquid	Kinks guidewire	Doesn't use scalpel	Forgets dilator	Inserts catheter without holding guidewire	Fails to confirm placement	Forgets guidewire after placing catheter	Doesn't advance to blue tape	Drop tool	Wrong step order	Other
Number of times event occurs	1														
	2														
	3														
	4														
	5														
	6														

Other Events / Comments Not Listed:

REFERENCES

1. Mansfield PF, Hohn DC, Fornage BD, Gregurich MA, DM. OTA. Complications and failures of subclavian-vein catheterization. *N Engl J Med*. 1994;331(26):1735-1738.
2. Taylor RW, Palagiri AV. Central venous catheterization. *Crit Care Med*. 2007;35(5):1390-1396.
3. Graham AS, Ozment C, Tegtmeyer K, Lai S, Braner D. Central venous catheterization. *N Engl J Med*. 2007;356(21):e21.
4. Merrer J, De Jonghe B, Golliot F, et al. Complications of femoral and subclavian venous catheterization in critically ill patients: a randomized control trial. *J Am Med Assoc*. 2001;286(6):700-707.
5. Sznajder JI, Zveibil FR, Bitterman H, Winer P, Bursztein S. Central vein catheterization failure and complication rates by three percutaneous approaches. *Arch Intern Med*. 1986;146(2):259-261.
6. Fares LG, Block PH, Feldman SD. Improved house staff results with subclavian cannulation. *Am Surg*. 1986;52:108-111.
7. Gualtieri E, Deppe SA, Sipperly ME, Thompson DR. Subclavian venous catheterization: greater success rate for less experienced operators using ultrasound guidance. *Crit Care Med*. 1995;23(4):692-697.
8. Scott DJ, Dunnington GL. The New ACS/APDS Skills Curriculum: moving the learning curve out of the operating room. *J Gastrointest Surg*. 2008;12(2):213-221.
9. Barsuk JH, Cohen ER, Feinglass J, McGaghie WC, Wayne DB. Use of simulation-based education to reduce catheter-related bloodstream infections. *Arch Intern Med*. 2009;169(15):1420-1423.
10. Kilbourne MJ, Bochicchio GV, Scalea T, Xiao Y. Avoiding technical errors in subclavian central venous catheter placement. *J Am Coll Surg*. 2008;208(1): 104-109.
11. Pugh C, Plactha S, Auyang E, Pryor A, Hungness E. Outcome measures for surgical simulators: is the focus on technical skills the best approach? *Surgery*. 2010;147(5):646-654.
12. Paul A. Harris, Robert Taylor, Robert Thielke, Jonathon Payne, Nathaniel Gonzalez, Jose G. Conde. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377-381.
13. IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.
14. ACS/APDS Surgery Resident Skills Curriculum. Avail able at: ⟨https://www.facs.org/education/program/apds-resident⟩ Accessed 19.03.16.
15. Bochiccchio GV, Joshi M, Bochicchio K, Nehman MS, Tracy JK, Scalea TM. Impact of obesity in the critically ill trauma patient: a prospective study. *J Am Coll Surg*. 2006;203(4):533-538.
16. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves OR performance: results of a randomized, double-blinded study. *Ann Surg*. 2002; 236(4):458-463.
17. Fried GM, Feldman LS, Vassiliou MC, et al. Proving the value of simulation in laparoscopic surgery. *Ann Surg*. 2004;240(3):518-525.
18. D'Angelo AD, Cohen ER, Kwan C, et al. Use of decision-based simulations to assess resident readiness for operative independence. *Am J Surg*. 2015;209(1): 132-139.

JOURNAL OF SURGICAL RESEARCH 199 (2015) 23–31

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.JournalofSurgicalResearch.com

Association for Academic Surgery

Residents' perception of skill decay during dedicated research time

Anne-Lise D. D'Angelo, MD, MS Ed, Rebecca D. Ray, PhD, Caitlin G. Jenewein, BS, Grace F. Jones, and Carla M. Pugh, MD, PhD*

Department of Surgery, School of Medicine and Public Health, University of Wisconsin, Madison, Wisconsin

ARTICLE INFO

Article history:
Received 3 January 2015
Received in revised form
9 June 2015
Accepted 17 June 2015
Available online 23 June 2015

Keywords:
Skill decay
Education
Resident
Simulation
Confidence
Self-efficacy

ABSTRACT

Background: Surgery residents may take years away from clinical responsibilities for dedicated research time. As part of a longitudinal project, the study aim was to investigate residents' perceptions of clinical skill reduction during dedicated research time. Our hypothesis was that residents would perceive a greater potential reduction in skill during research time for procedures they were less confident in performing.

Materials and methods: Surgical residents engaged in dedicated research training at multiple training programs participated in four simulated procedures: urinary catheterization, subclavian central line, bowel anastomosis, and laparoscopic ventral hernia (LVH) repair. Using preprocedure and postprocedure surveys, participants rated procedures for confidence and difficulty. Residents also indicated the perceived level of skills reduction for the four procedures as a result of time in the laboratory.

Results: Thirty-eight residents (55% female) completed the four clinical simulators. Participants had between 0–36 mo in a laboratory ($M = 9.29$ mo, standard deviation $= 9.38$). Preprocedure surveys noted lower confidence and higher perceived difficulty for performing the LVH repair followed by bowel anastomosis, central line insertion, and urinary catheterization ($P < 0.05$). Residents perceived the greatest reduction in bowel anastomosis and LVH repair skills compared with urinary catheterization and subclavian central line insertion ($P < 0.001$). Postprocedure surveys showed significant effects of the simulation scenarios on resident perception for urinary catheterization ($P < 0.05$) and LVH repair ($P < 0.05$).

Conclusions: Residents in this study expected greater skills decay for the procedures they had lower confidence performing and greater perceived difficulty. In addition, carefully adapted simulation scenarios had a significant effect on resident perception and may provide a mechanism for maintaining skills and keeping confidence grounded in experience.

1. Introduction

Concerns over resident readiness for operative independence have contributed to national discussions on the general surgery training curriculum [1–6]. Restructuring of the training curriculum [7], including changing the length of training [8], timing of specialization [8], and adding transition to care fellowships, is currently being debated [5,6].

* *Corresponding author*. Department of Surgery, University of Wisconsin, 600 Highland Ave., K6/135 CSC, Madison, WI 53792. Tel.: +1 608 262 0466; fax: +1 608 265 5963.

E-mail address: ad'angelo@uwhealth.org (A.-L.D. D'Angelo).

0022-4804/$ – see front matter © 2015 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jss.2015.06.040

Interestingly, there has been less discussion regarding the impact dedicated research time during residency has on surgical knowledge and skill acquisition. Residents that pursue research experiences in the middle of residency interrupt their clinical skills training before reaching mastery in certain skills. This may leave them vulnerable for skills decay and contribute to a lack of competency at the end of residency training.

General surgery training is unique in that many programs offer residents the opportunity to perform 1 y to 3 y of dedicated research time during the middle of residency training [9]. Almost 40% of residents training at Accreditation Council for Graduate Medical Education and National Resident Matching Program affiliated programs spend at least 1 y engaged in research [10]. The vast majority of residents start their dedicated research time after 2 y or 3 y of clinical training [10]. Although residents engaged in dedicated research training are significantly more likely to pursue fellowship training and less likely to hold private practice positions, little is known about how this research time affects operative performance [10,11].

Residents who pursue research experience may return to clinical practice with less surgical knowledge and skill than their peers who do not take time away for research. Factors known to contribute to skills decay include time away from task performance; level of knowledge mastery before time away and task characteristics [12,13]. There is a lack of research on skills decay during surgical research fellowships. Moreover, it is not known how skills decay during research fellowships impacts resident competency at the end of their training. Evaluating the potential reduction in surgical skill and knowledge during dedicated research time is critical to ensuring competent, independent performance at the end of training.

The aim of this study was to investigate residents' perceptions of clinical skill reduction during dedicated research time. We sought to evaluate the relationship between residents' perceived skill reduction, procedural confidence, and perceived task difficulty before and after performing simulation-based procedures. Our hypothesis was that residents would perceive a greater potential reduction in skill during research time for procedures they were less confident in performing. Our hypothesis is in line with other research noting that previously mastered skills are less prone to skills decay [12].

2. Methods

2.1. Setting and participants

Study participants ($n = 38$) were residents engaged in dedicated laboratory time from multiple general surgery training programs enrolled in a longitudinal simulation-based assessment study. This article evaluates data from the first data collection period of a longitudinal simulation-based assessment study. Residents who were currently in their clinical training were excluded from this study. Data collection occurred at five sites in three midwestern cities: Madison, WI; Chicago, IL; and Rochester, MN.

This study was approved by the University of Wisconsin Institutional Review Board and written informed consent was obtained from all participants.

2.2. Surveys

2.2.1. General survey

Before performing the simulated procedures, residents completed a general survey designed to collect information on demographics, years of general surgery training, years of dedicated laboratory work, and current call schedule. This survey also used a 5-point Likert scale to assess perceived reduction in global clinical and surgical skills and procedure-specific (urinary catheterization, subclavian central line insertion, bowel anastomosis, and laparoscopic ventral hernia [LVH] repair) performance during dedicated laboratory time.

2.2.2. Procedure related surveys

After the general survey, residents completed a preprocedure survey designed to assess confidence in and perceived difficulty in performing specific procedures steps and the entire surgical task. Survey items for each procedure are as follows: urinary catheterization (identify anatomy, problem solve difficulties with insertion, and successfully perform entire procedure); subclavian central line insertion (identify landmarks, cannulate the subclavian vein, and successfully perform entire procedure); bowel anastomosis (select correct suture, select correct stitch, and successfully perform entire procedure); and LVH repair (plan port placement, completely visualize the hernia defect, and successfully perform entire procedure). Confidence and perceived difficulty was assessed using a 5-point Likert scale (1 = Not confident; 5 = Extremely confident; and 1 = Not difficult; 5 = Extremely difficult, respectively). After completing each simulated procedure, residents completed a postprocedure survey designed to re-assess confidence and perceived difficulty of performing the surgical tasks. The same questions and 5-point Likert scales were used in the preprocedure survey and post-procedure survey.

2.3. Simulated procedures

After completing the general survey and preprocedure survey, participants completed the four simulated clinical procedures: urinary catheterization, subclavian central line insertion, bowel anastomosis, and LVH repair. These procedures were purposefully selected and designed to provide opportunities for both decision making and technical skill performance. Participants had 15 min to complete each simulated task with additional transition time between stations. This amount of time was selected to balance data collection requirements with participant fatigue. Simulator development was based on prior cognitive task analysis [14–16], and all stations were reviewed by experts before data collection. The focus of each station was predetermined to allow for participants to engage in major decisions and technical hurdles. The procedures were randomized with a Latin square. Before starting each station, researchers read an introductory narrative providing information on the simulator

and the station task. Next, participants were read a patient scenario before starting the procedure.

2.3.1. *Urinary catheterization*

This station was designed to represent four different patient scenarios (two female and two male) requiring urinary catheterization. Each participant was presented with three of the four possible patient scenarios in a randomized fashion. The patient scenarios included a 27-y old female trauma patient, 45-y old female operative patient, 74-y old male with rectal cancer, and 67-y old postoperative male with urinary retention. Patient scenarios were randomized by data collection site. The task in each scenario was to place a urinary catheter. Participants were informed that the perineum had already been prepped in the usual sterile fashion. Five urinary catheters (16F Foley, 16F Coude, 16F 3-way Foley, 12F Foley, and 10F Foley) were available for participants to use during this station along with lubricant and sterile water. Participants also had the option to place a urology consult. We theorized the urinary catheterization station would be easier from a technical skills aspect. This led us to include the choice of catheter in the simulation and a wide variation in clinical presentations.

2.3.2. *Subclavian central line insertion*

This station was designed to represent a hypotensive, tachycardic, and febrile patient in need of central venous access. The task was to place a subclavian central line. The simulator (CentraLineMan System; Simulab Corp, Seattle, WA) consisted of a realistic anatomic representation of a right upper adult torso with internal and external jugular and subclavian veins; carotid and subclavian arteries and venous and arterial “blood”. The simulator was placed in Trendelenburg position. Participants were informed that the patient had been prepped and draped in the usual sterile fashion and the skin had been injected with local anesthetic. A complete central line insertion kit with all the necessary equipment to perform a subclavian central line insertion, except an ultrasound machine, was provided. We theorized that residents would have varying levels of skill in performing this procedure. This led us to make a decision to allow the entire natural, built in complexity of central line placement to stand ‘as is’. We expected a wide variety of performance and completion rates in the allotted time. Safety in needle angle and maintenance of sterility are two examples of important decisions and possible errors.

2.3.3. *Bowel anastomosis*

This station was designed to represent a trauma patient in the operating room with an abdominal gunshot wound. The task was to inspect and repair the injured bowel. The simulator consisted of bovine intestine arranged on a tray with accompanying artificial blood and a single bullet. Two full thickness injuries in close proximity were located at the antimesenteric border [17]. One injury was larger (1×1 cm) with jagged edges and the other injury was smaller (0.5×0.5 cm). A trained researcher considered to be at the level of a medical student acted as a surgical assistant during the simulated procedure. Except electrocautery and stapler devices, all necessary open surgical instruments along with a selection of suture (3-0 Vicryl (Ethicon, Somerville, NJ), Nylon, Silk, PDS (polydioxanone; Ethicon), and Prolene(Ethicon)) were provided to complete the repair. Participants were required to decide how to repair the injury. We theorized that residents would have less experience independently completing this task. Choice of repair technique, suture selection, and stitch selection provided multiple decision points for participants to make.

2.3.4. *LVH repair*

This station was designed to represent an intraoperative patient with a ventral hernia. Most of the exposure had already taken place. The task was to perform a portion of the LVH repair by securing the mesh with the use of transfacial sutures and a tacker. The simulator is composed of a base covered with a simulated abdominal wall including a 10 cm $\times$ 10 cm ventral hernia [14–16]. Participants were informed that abdominal access had been achieved, the abdomen was insufflated, ports had been placed, and a piece of mesh inserted in the abdomen. The abdominal mesh was a 16×16 cm piece with four anchoring sutures already in place. Two of the four transfacial sutures had already been brought up through the abdominal wall and secured with hemostats. Participants were instructed to bring up the remaining transfacial sutures using a suture passing device and place the first five securing tacks. A tray of all the necessary open and laparoscopic instruments, except electrocautery, was available for use by the participants. A trained researcher considered to be at the level of a medical student acted as the operative assistant. We theorized that residents would have a wide range of actual experience as first assistant and operative surgeon at this step of the procedure. To assess port placement decisions, participants were required to diagram port location for laparoscopic hernia repairs on paper before starting the simulation.

Proficient use of laparoscopy and coordination with the suture passer device would be required to achieve adequate hernia coverage and well-placed mesh. These steps provided cognitive and technical complexity at this station. Ports were preplaced to provide participants with the maximal amount of time to perform mesh securing and tacking.

2.4. *Data analysis*

Demographic variables were averaged. Participant responses to preprocedure and postprocedure survey confidence and perceived difficulty ratings were averaged, and means were compared with repeated measures analysis of variance and paired *t*-tests. Pearson correlations were used to test associations between perceived skill reduction, total months spent in the laboratory, and preprocedural and postprocedural confidence and difficulty ratings.

3. Results

3.1. *Demographics*

Thirty-eight surgery residents, engaged in dedicated laboratory work (55% female), participated in the study. Residents had already completed 1 ($n = 1$), 2 ($n = 25$), or 3 ($n = 11$) clinical years at the time of participation in the study. One participant

did not indicate number of clinical years completed. Overall, residents had completed an average of 9.29 mo (standard deviation (SD) = 9.38) in the laboratory and planned to spend a mean of 16.32 additional months (SD = 8.40) doing research at the time of study participation.

Most residents (66%) had on-call responsibilities. Median on-call responsibilities was 1–2 shifts per month. A minority of residents (37%) were engaged in moonlighting work during their laboratory time. Those that performed moonlighting work had a median number of 3–4 shifts per month. Figure 1 depicts the percentage of residents that performed bedside or surgical procedures during on-call or moonlighting work.

3.2. *Perceived skill reduction*

Residents' perceived reduction in both global clinical and procedure-specific surgical skills is displayed in Figure 2. Among procedure-specific skills, residents perceived a greater reduction in bowel anastomosis ($M = 3.0$, $SD = 0.96$) and LVH repair ($M = 2.84$, $SD = 1.03$) skills compared with subclavian central line insertion ($M = 2.13$, $SD = 0.81$) and urinary catheterization ($M = 1.50$, $SD = 0.51$; $F(1,37) = 36.59$, $P < 0.001$; Fig. 2). The biggest perceived reduction in global clinical skills was in technical surgical skills ($M = 3.19$, $SD = 0.95$) and knowledge of procedure steps ($M = 3.00$, $SD = 0.76$, $F(6210) = 34.44$, $P < 0.001$, all pairwise comparisons $P < 0.01$). The months residents had already spent in the lab correlated with greater perceived reduction in technical surgical skills ($r(37) = 0.40$, $P = 0.007$) and knowledge of procedure steps ($r(37) = 0.36$, $P = 0.015$).

3.3. *Pre and post procedure confidence and difficulty ratings*

Participants reported the most preprocedure confidence for successfully performing the urinary catheterization procedure ($M = 4.24$ "very confident", $SD = 0.71$), followed by subclavian central line insertion ($M = 3.37$ "moderately confident", $SD = 0.91$), bowel anastomosis ($M = 2.40$ "somewhat confident", $SD = 1.08$), and LVH repair ($M = 2.13$ "somewhat confident", $SD = 1.02$; $F(3,111) = 37.22$ $P < 0.001$, all pairwise comparisons $P < 0.05$; Fig. 3). Similarly, participants predicted that the urinary catheterization task would be the least difficult ($M = 1.22$ "not difficult", $SD = 0.42$), followed by the subclavian central line insertion ($M = 2.11$ "somewhat difficult", $SD = 0.74$), bowel anastomosis ($M = 2.97$ "moderately difficult", $SD = 0.83$), and LVH repair ($M = 3.42$ "moderately difficult", $SD = 1.08$, $F(3,102) = 39.59$, $P < 0.001$, all pairwise comparisons $P < 0.005$, Fig. 4). These findings are consistent with our hypothesis regarding procedure experience.

Comparisons of preprocedure and postprocedure confidence are detailed in Figure 3A–D. After completing the simulated surgical tasks, participants reported significantly decreased confidence performing the entire urinary catheterization task ($M = 3.16$, $SD = 1.05$, $t(37) = 6.05$, $P < 0.001$) and specific procedure steps (identifying the relevant anatomy ($M = 3.92$, $SD = 0.82$, $t(37) = 3.47$, $P = 0.001$) and problem solving difficulties with urinary catheter insertion ($M = 3.24$, $SD = 0.88$, $t(37) = 5.28$, $P < 0.001$)). Participant confidence improved for completely visualizing the hernia defect ($M = 3.11$, $SD = 1.11$, $t(37) = -2.66$, $P = 0.012$) on the LVH task. All other average confidence ratings remain basically unchanged for the other simulated procedures.

Comparisons of preprocedure and postprocedure difficulty are detailed in Figure 4A–D. Similar to the confidence ratings, participants rated the urinary catheter task (identifying the anatomy, problem solving difficulties with insertion of the catheter and the successful completion of the entire task) as more difficult than predicted (all P values <0.01). Regarding central line, identifying the relevant landmarks was also rated as more difficult than expected ($M = 1.87$, $SD = 0.78$, $t(36) = -7.94$, $P < 0.001$). Conversely, after performing the simulation tasks, the LVH task (planning port placement, completely visualizing the hernia, and successfully performing the entire task) was rated as less difficult than previously predicted (all P values <0.05). Participants also rated selecting the correct suture for the bowel anastomosis as less difficult than expected ($M = 2.08$, $SD = 1.05$, $t(36) = 2.22$, $P = 0.033$).

Residents who reported engaging in surgical procedures during on-call responsibilities were more likely to have higher preprocedure and postprocedure confidence in the bowel anastomosis (pre $r(37) = 0.455$, $P = 0.004$; post $r(37) = 0.405$, $P = 0.012$); higher preprocedure and postprocedure confidence in conducting the entire LVH (pre $r(37) = 0.488$, $P = 0.002$; post $r(37) = 0.334$, $P = 0.041$). Engaging in surgical procedures also negatively correlated with difficulty ratings both preprocedure and postprocedure for the bowel anastomosis (bowel pre $r(37) = -0.390$, $P = 0.017$, post $r(37) = -0.497$, $P = 0.002$) and LVH repair (LVH pre $r(37) = -0.398$, $P = 0.016$, post $r(37) = -0.331$, $P = 0.042$). There were no significant correlations between

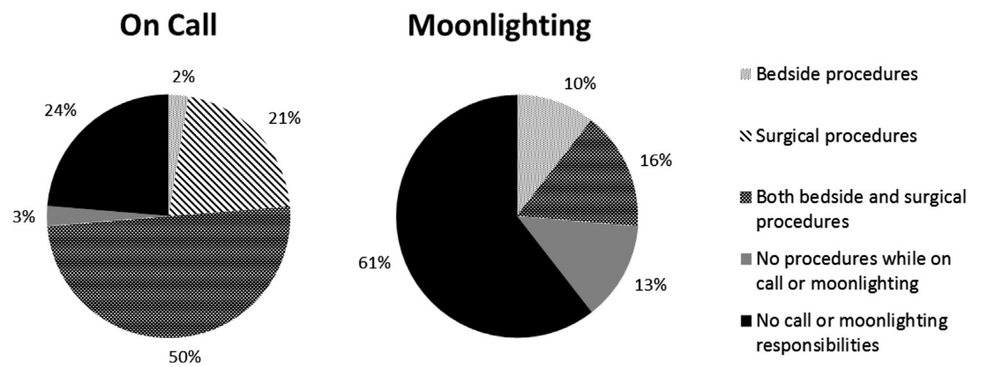


Fig. 1 – Proportion of residents engaged in procedures during on-call and moonlighting shifts.

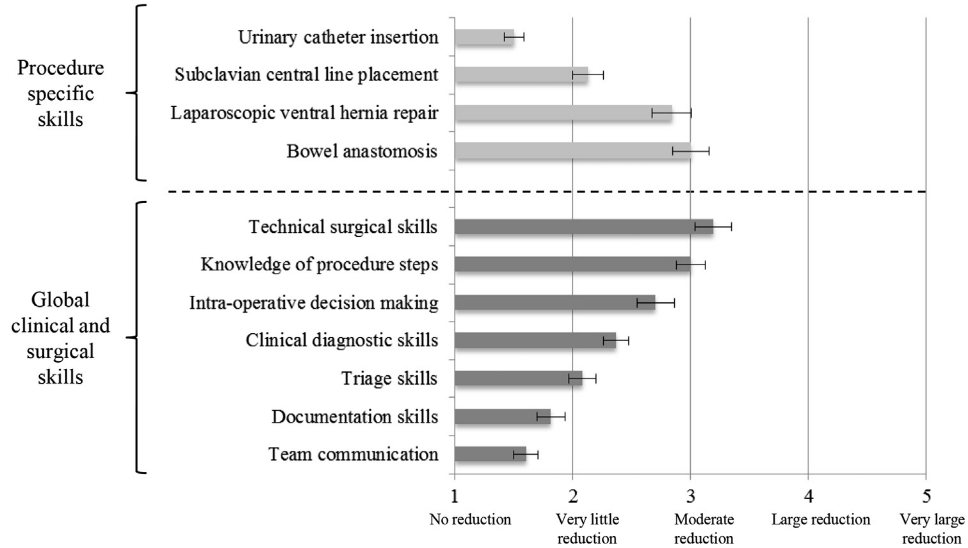


Fig. 2 – Average predicted reduction in procedure-specific and global clinical and surgical skills.

on-call surgical or bedside procedures and confidence or difficult performing the subclavian central line or urinary catheterization. There were no relationships between whether someone engaged in either bedside or surgical moonlighting procedures and confidence or difficulty ratings.

There were no significant correlations between months spent in the laboratory and perceived confidence and difficulty before and after simulation or changes in confidence and difficulty ratings before and after performing the simulation tasks.

3.4. *Preprocedure confidence and confidence change*

Confidence ratings before the simulation tasks correlated negatively with the amount of change in confidence in performing three of the four tasks (Table 1). Lower levels of confidence for the urinary catheterization, central line insertion, and bowel anastomosis predicted increased confidence in performing the entire procedure after the simulation (all *P* values <0.05).

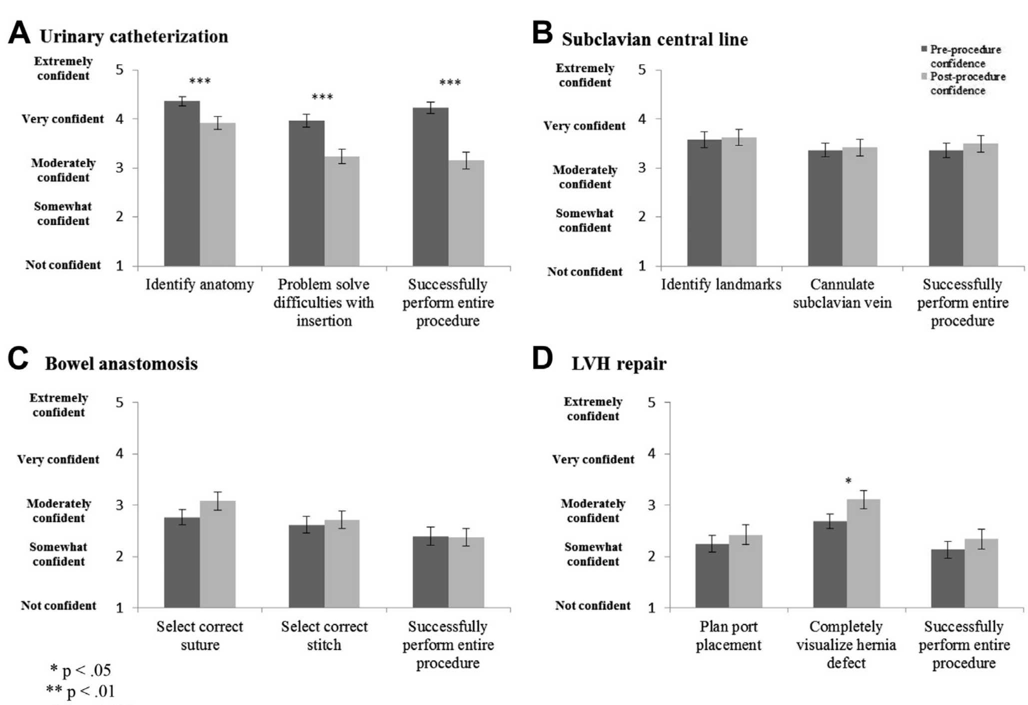


Fig. 3 – Average preprocedure and postprocedure confidence for four procedures (urinary catheterization [A], subclavian central line [B], bowel anastomosis [C], and LVH repair [D]).

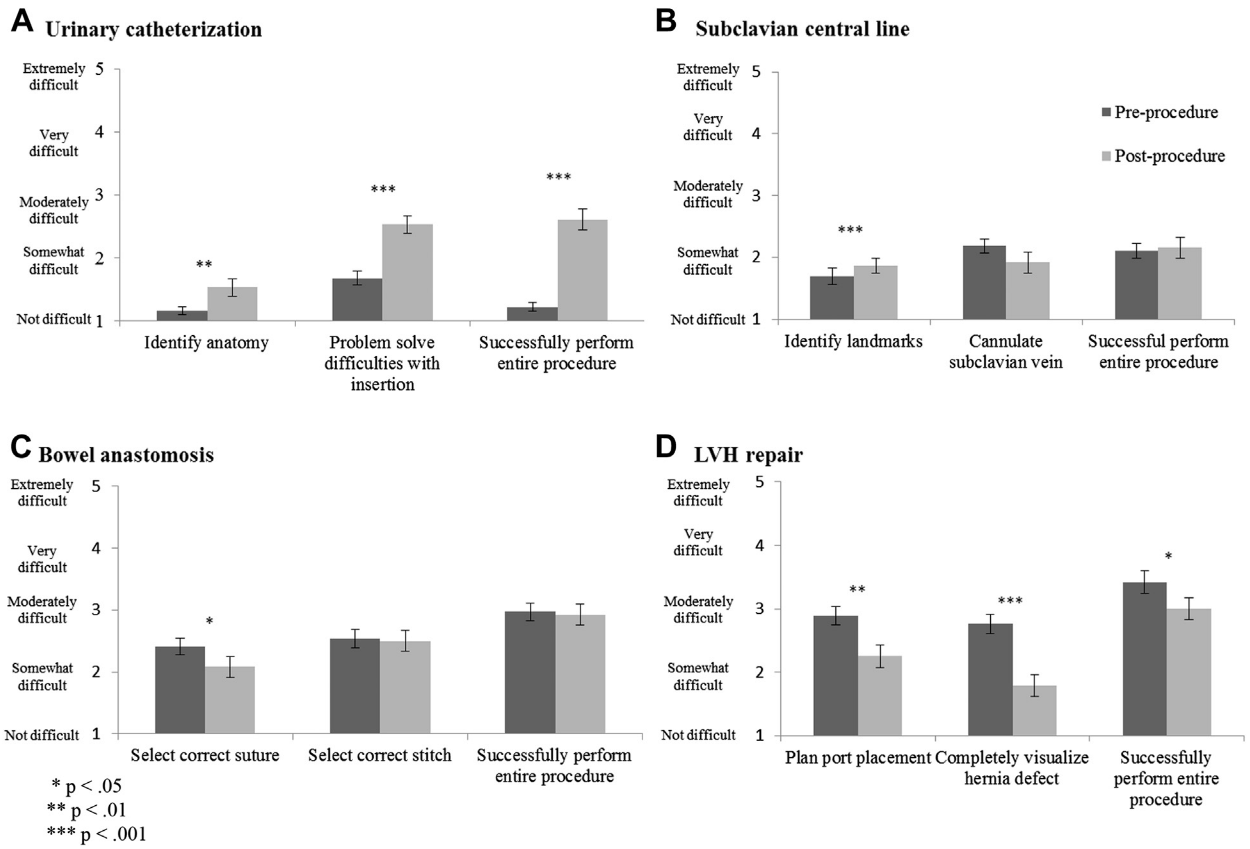


Fig. 4 – Average preprocedure and postprocedure perceived difficulty for four procedures (urinary catheterization [A], subclavian central line [B], bowel anastomosis [C], and LVH repair [D]).

3.5. *Perceived skill reduction and resident confidence and procedural difficulty*

Residents perceived greater skills reduction for those procedures that they had lower confidence in performing (Table 2). Specifically, predicted skill reduction in placing a subclavian central line negatively correlated with both before and after confidence ratings for performing the subclavian central line task ($P < 0.005$). Perceived skill reduction placing a urinary catheter correlated negatively with preprocedure confidence ratings only for the urinary catheter task ($P < 0.005$) and bowel anastomosis ($P < 0.005$). Predicted skill reduction performing a LVH repair was unrelated to either before simulation or after simulation confidence ratings for performing the LVH task.

The more skill reduction participants predicted in performing a subclavian central line insertion, the more difficult they rated completing the task both before and after performing the task (all P values <0.05, Table 2). Skill reduction ratings for the bowel anastomosis correlated with presimulation task difficulty ($r(37) = 0.34$, $P = 0.038$, Table 2), however, not with postsimulation task difficulty ratings. This can be interpreted to mean that the more bowel anastomosis skill reduction one predicted, the more difficult one perceived the entire task before the simulation. LVH skill reduction correlated with difficulty ratings of successfully performing the entire LVH task postsimulation ($r(37) = 0.32$, $P < 0.05$; Table 2).

4. Discussion

This study evaluated residents' perception of skill reduction during their time in the laboratory. Data were collected during the first time point of a longitudinal simulation-based skills assessment study. Participants in this study were residents at multiple general surgery training programs who were engaged in dedicated research fellowships. We found residents had varying perceptions of skill reduction during dedicated research fellowships based on procedure type. Additionally, performance of a simulated procedure differentially influenced their ratings of confidence and task difficulty based on procedure type and task focus. Tasks were purposefully modified to allow for a variety of task planning decisions and technical complexities.

Before performing the simulated procedures, residents perceived the greatest skill reduction for bowel anastomosis followed by LVH repair, subclavian central line insertion, and finally, urinary catheter insertion. Most residents in this study had completed two or three clinical years at the time of participation. Based on their level of training, these residents may not have reached mastery on the LVH repair or bowel anastomosis procedures. This is supported by our results demonstrating lower confidence and higher perceived difficulty for performing the LVH repair and bowel anastomosis than urinary catheterization or central line insertion. Prior work in skills decay has found that the amount of

Table 1 – Correlations between preprocedure confidence and the change in confidence before and after the simulated procedure.

Procedure	Correlation with preprocedure confidence	P-value
Urinary catheterization		
Identify relevant anatomy	−0.237	0.076
Problem solve difficulties inserting catheter	−0.368	0.011*
Successfully perform entire task	−0.388	0.008**
Subclavian central line		
Identify appropriate landmarks	−0.301	0.033*
Cannulate subclavian vein	−0.230	0.083
Successfully perform entire task	−0.299	0.034*
Bowel anastomosis		
Select correct suture	−0.225	0.090
Select correct stitch	−0.305	0.031*
Successfully perform entire task	−0.470	0.001***
Laparoscopic ventral hernia repair		
Plan proper port location	−0.227	0.085
Completely visualize hernia defect	−0.193	0.123
Successfully perform entire task	−0.139	0.203

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

overlearning (or mastery) is a critical determinant of skill and knowledge retention [12,18]. Overlearning is training beyond that required for proficiency. Residents that have not had opportunities for overlearning on certain tasks such as the bowel anastomosis or LVH repair would be expected to have lower confidence and perceive higher risks of skills decay. Further work is needed to determine if their perception has any effect on performance.

Residents who had on-call responsibilities and engaged in surgical procedures reported higher preprocedure and postprocedure confidence for performing the bowel anastomosis and LVH repair. This indicates that exposure to surgical procedures during dedicated research time may allow for maintenance of self-efficacy and potentially maintenance of proficiency. There were no correlations between engagement in on-call surgical procedures and urinary catheterization and subclavian line confidence or difficulty ratings. This may relate to the fact that residents were overall more confident in performing these procedures and expected less skill decay. Periodic exposure to surgical tasks decreases the retention interval and may support proficiency maintenance during time away from training [12,19].

Residents also believed that they would have the greatest reduction in global technical surgical skill and knowledge of procedure steps. This perceived skill reduction was correlated with time already spent in the laboratory. Compared with clinical diagnostic, triage, documentation, and team communication skills, residents may have less experience with technical surgical skills and knowledge of procedure steps by the end of their second or third clinical year. This is consistent with the overlearning theory of skills decay. If residents have not had the opportunity for overlearning, global skills and procedure-specific skills may be at risk for skills decay during dedicated research fellowships [12].

In the absence of any feedback on their performance, participants' preprocedure and postprocedure confidence and difficult ratings changed according to actual task complexity. For example, the urinary catheter procedure was purposefully made complex. Before performing the procedure, participants rated the urinary catheterization as having no to very little skills reduction along with high confidence and low difficulty ratings. However, as expected, there were significant reductions in confidence and increases in difficulty ratings for all three items related to urinary catheter insertion (identifying the anatomy, problem solving difficulties with insertion of the catheter, and the successful completion of the entire task). Conversely, the LVH task was shortened and made to focus on a small part of the procedure. Before performing the procedure, participants rated the LVH repair as having moderate skills reduction, somewhat to moderate confidence ratings and moderately to very difficult ratings. As expected, after performing the LVH simulated procedure, there were significant increases in confidence and decreases in difficulty ratings. This indicates that specific modifications of performing a simulated task can influence ratings of procedure confidence and difficulty. Residents may become more confident on tasks that were easier than expected and become less confident on tasks that were more difficult than expected. Prior work has shown that during periods of inactivity, self-efficacy can be maintained even though skill performance declines [20]. Participation in periodic skills evaluations combined with simulation-based training can be useful to maintain the relationship between self-efficacy and actual performance. Additionally, it is important to be cognizant of simulation scenario design for decision making and technical skills as it may impact residents' self-perceptions of skill even in the absence of feedback.

Table 2 – Correlations of perceived skill reduction and preprocedure and postprocedure confidence and difficulty ratings for four procedures.

Perceived skill reduction	Confidence		Difficulty	
	Preprocedure	Postprocedure	Preprocedure	Postprocedure
Urinary catheterization	−0.560***	−0.051	0.117	−0.228
Subclavian central line	−0.578***	−0.582***	0.346*	0.368*
Bowel anastomosis	−0.496**	−0.275	0.343*	0.134
Laparoscopic ventral hernia repair	−0.186	−0.065	−0.023	0.321*

Study limitations largely relate to the simulated tasks and scenarios. Residents in this study performed four simulations that were chosen to represent a range of expected mastery for their skill level. However, these procedures are a small subset of possible clinical scenarios that a resident might see in surgical practice. As such, our results may not be generalizable across all procedures and expectations for mastery. Furthermore, the simulation exercises were designed to be conducted within a specific time frame. This may have influenced how residents' performed the task and possibly affected the difficulty of the procedure. However, we purposefully designed the simulations to contain opportunities for both decision making and technical performance in the amount of time provided. This study focuses on self-reported measures that may not correspond with objective performance. Residents' perception of their performance may be different than their actual performance. In addition, feedback was not provided before completion of the postsurveys. Finally, we may not have found significant relationships between moonlighting surgical procedures and confidence or difficulty ratings given our small sample size, which limits statistical power.

This study showed that residents expect moderate skills decay on more difficult tasks (bowel anastomosis and LVH repair) and for technical surgical skills and knowledge of procedure steps. Despite this expectation for skills reduction during dedicated research fellowships, few programs have focused clinical skills curricula to reduce skills decay or remediate performance. Prior research has demonstrated that training to mastery has a protective effect for skills decay [21]. Those skills that residents have not trained to mastery may be more vulnerable to skills decay and require additional work for remediation. Incorporation of these activities in resident training may inform training interventions that can be used throughout dedicated research time. Our results show that simulation-based assessments and training affect resident perception of skill and may provide a mechanism for maintaining skills [18] and keeping confidence grounded in performance.

Acknowledgment

The authors thank Patrick Barlow, PhD, for his contribution to survey design and data collection.

Funding for this study came from the Department of Defense grant W81XWH-13-1-0080 entitled, "Psychomotor and Error Enabled Simulations: Modeling Vulnerable Skills in the Pre-Mastery Phase" and National Institutes of Health grant 1F32EB017084-01 entitled "Automated Performance Assessment System: A New Era in Surgical Skills Assessment".

Authors' contributions: A.-L.D.D., R.D.R., and C.M.P. contributed to study conception and design. A.-L.D.D., R.D.R., C.G.J., G.F.J., and C.M.P. contributed to study data collection. A.-L.D.D., R.D.R., and C.M.P. did the analysis and interpretation of data. A.-L.D.D., R.D.R., C.G.J., G.F.J., and C.M.P. contributed in writing the article and critical revision of the article.

Disclosure

The authors reported no proprietary or commercial interest in any product mentioned or concept discussed in this article.

REFERENCES

[1] Coleman JJ, Esposito TJ, Rozycki GS, et al. Early subspecialization and perceived competence in surgical training: are residents ready? J Am Coll Surg 2013;216:764.

[2] Nakayama DK, Taylor SM. SESC Practice Committee survey: surgical practice in the duty-hour restriction era. Am Surg 2013;79:711.

[3] Bucholz EM, Sue GR, Yeo H, Roman SA, Bell RH Jr, Sosa JA. Our trainees' confidence: results from a national survey of 4136 US general surgery residents. Arch Surg 2011;146:907.

[4] Mattar SG, Alseidi AA, Jones DB, et al. General surgery residency inadequately prepares trainees for fellowship: results of a survey of fellowship program directors. Ann Surg 2013;258:440.

[5] Richardson JD. ACS transition to practice program offers residents additional opportunities to hone skills. The Bulletin of the American College of Surgeons. Available from: http://bulletin.facs.org/2013/09/acs-transition-to-practice-program-offers-residents-additional-opportunities-to-hone-skills/. Published 2013. Accessed March 6th, 2014.

[6] Hoyt BD. Executive Director's report: looking forward - February 2013. Chicago, IL: Bulletin of the American College of Surgeons. Available from: http://bulletin.facs.org/2013/02/looking-forward-february-2013/; 2013. Published 2014. Accessed March 6th, 2014.

[7] Lewis FR, Klingensmith ME. Issues in general surgery residency training–2012. Ann Surg 2012;256:553.

[8] Klingensmith ME, Lewis FR. General surgery residency training issues. Adv Surg 2013;47:251.

[9] Roberston CM, Klingensmith ME, Coopersmith CM. Long-term outcomes of performing a postdoctoral research fellowship during general surgery residency. Ann Surg 2007;245:516.

[10] Roberston CM, Klingensmith ME, Coopersmith CM. Prevalence and cost of full-time research fellowships during general surgery residency. Ann Surg 2009;249:155.

[11] Ellis MC, Dhungel B, Weerasinghe R, Vetto JT, Deveney K. Trends in research time, fellowship training, and practice patterns among general surgery graduates. J Surg Education 2011;68:309.

[12] Arthur W Jr, Bennett W Jr, Stanush PL, McNelly TL. Factors that influence skill decay and retention: a qualitative review and analysis. Hum Perform 1998;11:57.

[13] Schendel JD, Shields JL, Katz MS. Retention of motor skills: review. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences; 1978. p. 1. Tech. Paper 313.

[14] Pugh CM, DaRosa DA, Santacaterina S, Clark RE. Faculty evaluation of simulation-based modules for assessment of intraoperative decision making. Surgery 2011;149:534.

[15] Pugh C, Plachta S, Auyang E, Pryor A, Hungness E. Outcome measures for surgical simulators: is the focus on technical skills the best approach? Surgery 2010;147:646.

[16] Pugh CM, Santacaterina S, DaRosa DA, Clark RE. Intraoperative decision making: more than meets the eye. J Biomed Inform 2011;44:486.

[17] D'Angelo AD, Cohen ER, Kwan C, et al. Use of decision-based simulations to assess resident readiness for operative independence. Am J Surg 2015;209:132.

[18] Baldwin TT, Ford JK. Transfer of training: a review and directions for future research. Personnel Psychol 1988;41:63.
[19] Stefanidis D, Korndorffer JR, Markley S, Sierra R, Scott D. Proficiency maintenance: impact on ongoing simulator training on laparoscopic skill retention. J Am Coll Surg 2005; 202:599.
[20] Youngqiust ST, Henderson DP, Gausche-Hill M, Goodrich SM, Poore PD, Lewis RJ. Paramedic self-efficacy and skill retention in pediatric airway management. Acad Emerg Med 2008;15:1295.
[21] Hagman JD, Rose AM. Retention of military tasks: a review. Hum Factors 1983;25:199.